9-10-2018

Essays on Matching Markets: Course Allocation, Team Formation, and P2P Lending Markets

Hoda Atef Yekta
University of Connecticut - Storrs, hoda.atef_yekta@uconn.edu

Essays on Matching Markets:

Course Allocation, Team Formation, and P2P Lending Markets

Hoda Atef Yekta

University of Connecticut

Abstract

In this dissertation we analyze matching markets using two broad methodologies. In the first two of three essays, we use integer programming to design new practical markets, solving difficult organizational problems in which items like a seat in a classroom course or membership on a project team are allocated to agents who express their preferences to a centralized planner. In the third essay, we employ empirical modeling to study the signals sent among three types of agents in a peer-to-peer (P2P) lending market: borrowers, lenders, and a (rating-service) platform. The first two essays use extensive simulations, while the third uses statistical analysis on a large empirical dataset: four years of loan application and payment history from a prominent online P2P lending platform.

Our findings in the first essay show that large but manageable course allocation problems can be solved with various multi-stage optimization algorithms, providing much better outcomes than existing benchmarks from the literature on metrics of efficiency, fairness, and incentive-compatibility. We demonstrate robustness of our techniques by simulating over a variety of market parameters, including varying degrees of manipulation over a range of common to private value utility functions. These results show promising new practical designs that can satisfy more organizational objectives than previous methods.

In the second essay, we find that the much harder (interpersonal) quadratic-interaction optimization in an agent-based team formation game requires advanced computational technique. In the pursuit of a balance between the competing objectives efficiency and group stability, we explore the cutting-edge of computational operations research. In contrast to existing draft-based systems that favor distributed, intuitive heuristics, we solve the extremely difficult centralized

optimization problems directly, with positive results using a new technique for bi-level optimization in this context, based on a customized column generation framework. Again, the results produce a promising new optimization framework for a practical problem, and our results compare favorably with existing methods.

Finally, the results of the third essay show just how complex behavior in real-world matching markets can be through the empirical analysis of a real market. We verify a few intuitive results, but also find some counterintuitive interactions, in which a monotonically increasing signal from the market platform results in a non-monotonic return on investment (ROI), for example. Overall, we find risk-seeking behavior among peer investors that does not tend to pay off, and some strange disconnects in which, for example, investors favor "Debt Consolidation" loans despite inferior ROI, and have a prejudice against "Business" loans despite no significant evidence of poor performance.

Essays on Matching Markets:

Course Allocation, Team Formation, and P2P Lending Markets

Hoda Atef Yekta

B.S. Amirkabir University of Technology

M.S. Sharif University of Technology

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2018

APPROVAL PAGE

Doctor of Philosophy Dissertation

Essays on Matching Markets:

Course Allocation, Team Formation, and P2P Lending Markets

Presented by

Hoda Atef Yekta, B.S., M.S.

Major Advisor____

Dr. Robert Day

Associate Advisor____

Dr. David Bergman

Associate Advisor____

Dr. Mike Shor

University of Connecticut
2018

Acknowledgements

My graduate education has now reached the end of a long and wonderful journey. On this occasion, I would like to take this opportunity to thank individuals who helped me along the way.

First and foremost, I am honored to express my gratitude to my major advisor, Dr. Robert Day for his patience, guidance, support, and encouragement over the past five years. Without his advice and persistent help, this work would not have been possible.

I would like to extend my appreciation to others who have advised my studies, Dr. David Bergman, Dr. Ram Gopal, and Dr. Mike Shor. I am grateful for their invaluable impact on my education through their courses, suggestions, and advice. I am also thankful for the personal and professional support of many faculty members at Operations and Information Management Department at UCONN.

My deepest gratitude and love go to my kind parents and sister in Iran for their great support, kindness and encouraging through this long journey. Last, but certainly not least, I would like to express the deepest appreciation to my beloved husband, Hamed, for his encouragement, advice and kind supports in all steps of our life, he stood with me through good times and bad times, and our little son Hossein who was with us in the last steps of this journey.

Contents

Chapter 1. Introduction

Matching theory has a long history as a branch of game theory, and was recognized with the Nobel Memorial Prize in Economics in 2012. In recent decades, applications of matching theory and the framework of “marriage” and “roommate” markets have included successful programs for matching students to schools, applicants to public housing, workers to jobs, and courses to students, to name a few. In first two essays of this dissertation, we analyze new extensions of matching theory and simulate new applications. In a third essay, we turn our attention from non-monetary markets to an empirical study of peer-to-peer (P2P) lending markets, which often take the form of matching markets with a prescribed monetary contract (i.e., an exogenous platform-determined interest rate).

In the first essay of this dissertation, Chapter 2, we focus on a many-to-many matching problem in which a set of objects are distributed among a set of agents based on agent preference. A familiar example of this problem is the “course allocation problem” in which scarce classroom capacity is assigned to students when several particular courses experience significant excess demand. In recent years, several universities have adopted an algorithmic approach to the allocation of seats in courses, in which students place bids, and then seats are awarded according to a predetermined procedure or mechanism. Although designing the appropriate mechanism for translating bids into student schedules has received attention in the literature, there is currently no consensus on the best mechanism in practice. Thus, we introduce five new algorithms for this problem, using various combinations of matching algorithms, second-price concepts, and optimization, and compare our new methods to the natural benchmarks from the literature.

In a second essay, Chapter 3, we worked on an extension of the roommate problem, called the team formation problem (TFP), in which a group of N agents are to be partitioned into M groups of size N/M (or floor(N/M) and ceiling(N/M) when N/M is non-integer). Such problems arise in project management, military platooning, sports-league management, and in assigning MBA group-projects, for example. This problem can be modeled similar to the marriage or roommate problems in which we prefer "stable solutions," i.e., solutions in which no group of people would be mutually better off leaving their team assignment to form their own team. Unlike the marriage problem, however, there may be no stable solution, and a tractable algorithm for finding stable solutions when they exist is currently unknown. Further, partitioning the market in order to maximize total utility is an NP-hard problem. In this essay, we present a column generation scheme which provides orders of magnitude speedups over existing algorithms for partitioning the market with respect to a quadratic objective function. We also define new metrics for measuring stability and compare our solutions with those obtained by existing algorithms on measures of efficiency, fairness, stability, and the effect of strategic behavior.

In a third essay, Chapter 4, we turn our attention to peer-to-peer lending markets that matches many lenders to many borrowers. In these markets, lenders select a portfolio of unsecured personal loans to fund, in full or partially. Having collected a large dataset of publicly available loan information for over four years of loan origination requests (with all follow-up data through the completion of 36-month loan terms) from an anonymous lending platform, this study seeks to shed light on the interplay between the players in these markets, showing how signals from one class of participant effects the behavior of others using data analytics. In particular, we first explore the borrowers' disclosed personal information and analyze the response of the platform as well as the investors to these signals. Then, we analyze the response of investors to the platform signals and examine how closely investors follow the signals provided by the platform. Finally, we study the efficiency of both the borrower self-reported information as well as the P2P lending platform signals in predicting the success of each loan.

Chapter 2. Optimization-based Mechanisms for the Course Allocation Problem

Abstract:

In recent years, several universities have adopted an algorithmic approach to the allocation of seats in courses, in which students place bids (typically by ordering or scoring desirable courses), and then seats are awarded according to a predetermined procedure or mechanism. Designing the appropriate mechanism for translating bids into student schedules has received attention in the literature, but there is currently no consensus on the best mechanism in practice. In this paper, we introduce five new algorithms for this course allocation problem, using various combinations of matching algorithms, second-price concepts, and optimization, and compare our new methods to the natural benchmarks from the literature: the (proxy) draft mechanism and the (greedy) bidding-point mechanism. Using simulation, we compare the algorithms on metrics of fairness, efficiency, and incentive compatibility, measuring their ability to encourage truth-telling among boundedly-rational agents. We find good results for all of our methods, and that a two-stage full-market optimization performs best in measures of fairness and efficiency, but with slightly worse incentives to act strategically compared to the best of the mechanisms. We also find generally negative results for the bidding-point mechanism, which performs poorly in all categories. These results can help guide the decision of selecting a mechanism for course allocation or for similar assignment problems, such as project team assignments, or sports drafts, for example, where efficiency and fairness are of utmost importance, but incentives must also be considered. Additional robustness checks and comparisons are provided in a series of online appendices.

1. Introduction

The allocation of scarce classroom capacity to students is a challenging process for many universities when several particular courses experience significant excess demand. Similarities in students' preferences, limits on the number of semesters to complete a program, and a limited number of seats (especially in MBA programs) intensify the importance of what has become known as the *course allocation problem* (CAP). In the last three decades, several universities, especially highly-ranked (and thus heavily demanded) business schools like Stanford Business School, Harvard Business School, Ross School of Business at University of Michigan, and Columbia Business School have tried to improve their mechanisms for allocating courses among students and have tested different methods to manage and solve this problem (Krishna & Ünver, 2008).

In this paper, we consider several metrics relevant to a university in making these decisions, in order to compare the practicality of newly introduced optimization-based approaches and two standard benchmarks on simulated data. These metrics can be categorized into three groups and described as follows:

Measures of Efficiency

An allocation is *Pareto efficient* if no student can be made better off without making another student worse off. Some mechanisms are not even Pareto efficient with respect to the submitted preferences (i.e., may not be Pareto efficient even when all players reveal their preferences truthfully.)

The *allocative efficiency* of a market outcome measures the total market utility (satisfaction) over all players, reflecting that the goal of a university is to satisfy students as much as possible with their most preferred courses. Total market utility can be measured with a (traditional) cardinal utility objective function, or an ordinal utility objective (comparing only the relative rankings of courses), or a binary utility objective (simply measuring the total number of seats assigned by the market.)

Measures of Fairness

The range and standard deviation of individual student utility provide measures of dispersion among the students, with high values indicating outcomes that would be viewed as highly unequal among students and thus unfair. As with allocative efficiency, fairness can be measured using cardinal, ordinal, or binary utility measures. For example, under binary utility, the range and standard deviation measure differences in the total number of courses awarded to a student. Significant dispersion in the number of courses assigned to a student is problematic, given the basic expectation of getting a full or nearly full schedule of courses. Dispersion when using the other utility measures is similarly problematic.

Measures of Incentive Compatibility

To test the relative truth-inducing properties of the various course allocation mechanisms, we adapt a common-value utility approach from Kominers et al. (2010), in which a student's true utility for each course is simulated as the mean of a private-value component and a common-value component. This setup creates a high degree of value correlation and a natural method to behave strategically under highly uncertain conditions: moving the weighting factor from the private signal toward the common-value signal tends to increase one's chances of getting highly popular courses. We implement these boundedly-rational strategies using a range of weights (more and less extreme strategies) and among a higher and lower percentage of such "strategic" students. We can then measure the benefit or lack of benefit to honest and strategic students under a variety of scenarios, measuring where strategic play is most beneficial to those who employ such tactics, and most hurtful to those who remain honest.

Focusing on these metrics, our results show a great deal of promise for mechanisms that use an integer-programming approach. Our five new algorithms represent new combinations of existing ideas from the literature, including round-by-round allocation as in draft mechanisms, top-trading cycles from matching theory, second-price concepts from auction theory, and the use of optimization, both within rounds and globally across an entire market.

Round-by round allocation has the potential to increase fairness by ensuring that no student gets allocated her $t + 1^{\text{st}}$ course before all students have been allocated t courses (except when a student has exhausted her list of acceptable courses and thus can be allocated no more.) Therefore four of our five new algorithm variations investigate the use of round-by-round allocation for its potential to have good fairness properties.

Within the round-by-round framework, we pursue two different general approaches. The first emulates the top-trading-cycle (TTC) algorithm commonly discussed alongside the deferred-acceptance algorithm in matching theory, adapted here to be iterated for the multi-round allocation of several courses, as opposed to matching theory's typical unit-demand setting. This algorithm will be referred to as TTC for its connection to that classical approach.

Our second approach emulates auction theory in its use of second prices to squelch strategic manipulability. Though a Vickrey-Clarke-Groves (VCG) mechanism can be used to provide incentive compatibility in weakly-dominant strategies even for the general combinatorial auction setting, it uses actual currency and a payment scheme to control incentives. For universities, however, course allocation cannot be settled through monetary bidding and payments (a situation known in the literature as non-transferable utility, NTU) making the control of incentives more difficult. Still, although a bidding student has no value for unused points after participation in the market, a multi-round setting can potentially set up a situation in which unused bidding points in one round can be useful as points to be used in subsequent rounds, potentially returning some currency value and auction-like properties. We therefore investigate course allocation as a sequence of second-price auctions using bidding points in each round. This algorithm will be referred to as SP for second-price.

Each of these multi-round algorithms is also implemented using within-round optimization, creating optimization versions of each, our third and fourth algorithm variations, TTC-O and SP-O, respectively. These within-round optimizations tend to improve efficiency in harmony with the multi-round formats' focus on fairness, but the effect of this combined approach on incentives is not known, leading us to investigate the impact via simulation.

Our fifth algorithm attempts optimization over the entire market, first maximizing ordinal utility, followed by a second optimization to maximize cardinal utility subject to ordinal-utility optimality. This two-stage approach reflects a desire to avoid interpersonal cardinal utility comparisons as much as possible, instead reflecting a university's desire to have a large number of highly ranked course assignments. The magnitudes of bid-point revelations are instead used only for tie breaking, guaranteeing that any cardinal-utility comparison is based on intrapersonal effects, not just direct interpersonal comparisons, which are harder to justify. Our findings show that this approach of ordinal-then-cardinal optimization (which we abbreviated OC) performs quite well in the areas of efficiency and fairness.

The remainder of the paper is organized as follow. First, a comprehensive review of the current body of knowledge is discussed in §2. Next, we introduce our five new algorithm variations for solving the course allocation problem in §3. Then in §4 and §5, we discuss the results of two extensive simulations comparing these new algorithms to the most commonly discussed methods in the literature (the draft mechanism and BPM) using the measures of efficiency, fairness, and incentive compatibility discussed above. Concluding remarks are provided in §6.

Supporting material is provided in a set of appendices. Appendix A describes the simulation details for the efficiency/fairness experiments in greater depth. Additional results on incentive compatibility are described in Appendix B, and a brief additional discussion of a more recent alternative mechanism (A-CEEI) due to Budish (2011) is given in Appendix C. In Appendix D, we explore robustness under a more random setup while varying the amount of correlation among student preferences, finding that the fairness of our round-by-round optimization-based methods are more robust to extreme preference correlation, as might be expected.

2. Background and Literature Review

A review of the literature reveals five prominent alternative algorithm variations for CAP: the bidding-point mechanism, the draft mechanism (including a "proxy bidding" version), random serial dictatorship, the "Gale-Shapley Pareto-dominant market" mechanism, and the "Approximate Competitive

Equilibrium for Equal Income (A-CEEI)" mechanism. Each of these algorithms has its own strengths and weaknesses, and in the rest of this section, we mainly discuss the details of each. Moreover, at the end of this section, we also briefly review the other more recent papers for solving CAP.

Two main mechanisms have been most widely used by different universities, the *bidding-point mechanism (BPM)* and the *draft mechanism*. In BPM, a fixed bidding budget of points is assigned to each student, and a bid is submitted that allocates these points among her various courses of interest. Following a complete submission from all students, all bids for all courses are sorted into a single list, from highest to lowest, and accepted in turn, if eligible, one at a time. Each bid is considered if and only if the bidding student has not filled her schedule, this course does not conflict with her current assigned courses, and the course has not filled all its seats (Sönmez and Ünver, 2010). This algorithm (with small variations) has been used by several schools, including University of Michigan's Ross School, Kellogg Graduate School of Management at Northwestern, Johnson Graduate School of Management at Cornell, Columbia Business School, Haas School of Business at UC Berkeley, and Yale School of Management (Krishna & Ünver, 2008).

Budish (2011) describes the main issue of BPM, its unfair results, in which one student may receive no course for one semester and her unspent points are wasted, while another student may receive all of her desired courses, a result which unexpectedly happened frequently at University of Chicago's Booth School of Business. Indeed, BPM focuses only on maximizing allocative efficiency and does not consider equity among students. In fact, it compromises both equity in number of courses and fairness in value of courses among students in order to achieve a solution which greedily maximizes the total bid of assigned courses. Moreover, moving beyond a static, one-shot investigation for some highly popular courses, course prices may become distorted and chaotic if taken as signals for bidding from semester to semester. That is, the course price may become as high as the whole budget of a student in one semester, discouraging students in the next semester from bidding on it, only to see its price drop, leading demand to rise and fall sporadically over time.

In the draft mechanism, used by Harvard Business School beginning in the mid-1990s, a computer takes students' preferences over individual courses. Then, student proxies take turns in a "draft order," with one available course seat assigned to each student in each round. In the draft order procedure, students are first randomly ordered, and in even-numbered rounds their order is reversed (Budish & Cantillon, 2011).

The draft mechanism is fairer than BPM because the round-by-round allocation naturally decreases the range of the number of assigned courses per student. However, the initial random favoritism among students may result in successive bad luck for a student and causes dissatisfaction based on inequality of ordinal or cardinal utility. In fact, the draft mechanism only collects the student ordinal preferences and therefore cannot distinguish between a slight difference and a significant difference among preferences. Hence, a student who is almost indifferent between two courses C1 and C2 and has put C1 first may get this course while another student who wants C1 much more than C2 may lose it. Though a small decrease in utility of the first student could allow a large increase for the second, this opportunity to improve overall cardinal utility is lost by the draft mechanism.

Furthermore, in both BPM and the draft mechanism each student's belief about others' favorite courses affects her declared preferences, encouraging her to behave strategically and misreport her preferences (Budish & Cantillon, 2011). In fact, our simulations show that BPM is seriously vulnerable to student strategic behavior, with an intentional misreporting of preferences easily leading to a more favorable outcome.

Kominers et al. (2010) introduced a variation of the draft mechanism based on proxy-bidding, which showed improved performance with respect to incentives in their simulations, i.e., showing less benefit to those who choose to bid strategically. The approach is to generate true-bid values for simulated students based on an equally weighted average of a common-value and a personal value for each course, and moving the weight towards the common (popularity-based) value to simulate strategic behavior. We adapt this approach to the current setting in the experiments described in §4.3, as a first approximation of strategic behavior in a setting of bounded rationality (incomplete and uncertain information with no ability to devise a stochastically optimal strategy) as in their paper. Though their results show some improvement of incentives over the draft mechanism, it suffers some of the same problems, in which a commitment to a random prioritization of students results in lost opportunities for efficiency improvements as mentioned above. Also, some of their results are based on the assumption of non-overlapping courses, though our current study takes the more realistic assumption that some course sections overlap in time, and that different sections of the same course may be available, though only one can be taken.

A third algorithm discussed in the literature is the random serial dictatorship, in which students are sorted randomly and in each turn one student picks her entire bundle of courses from any seats still available. This mechanism is the only incentive-compatible mechanism for CAP in the literature (see Papai, 2001, Ehlers and Klause, 2003, and Hatfield, 2009) because student strategic behavior cannot change their assigned courses. However, fairness and equity are totally neglected in this method (Budish & Cantillon, 2011). Indeed, randomness of this sort is in favor of incentive compatibility in general: basing "prioritization" on declared preferences (rather than randomness) in general introduces opportunities for manipulation of the declared preferences. But given the quite bad fairness and efficiency for these a totally random prioritization, the perspective of the current paper is to find mechanisms which do not perform very badly on incentives, while offering drastic improvements on fairness and efficiency.

A fourth algorithm is the Gale-Shapley Pareto-dominant market mechanism (GS mechanism) which is developed by Sönmez and Ünver (2010) as a hybrid between BPM and the familiar GS deferred-acceptance algorithm for matching. This algorithm gives a fixed bidding budget to each student. Then, each student submits both a rank ordered list and a bid list for her desired courses, where these two lists need not agree with each other. (If the rank ordered lists and the bid lists agree with each other, GS and BPM result in the same solution.) In this version of the GS algorithm, first each student (proxy) points to her most preferred course based on rank and then proposes a match. Then, each course sorts all the offers and keeps the highest bids, up to capacity, and rejects the remaining ones. Students with rejected offers point to their next most preferred course, proceeding as in a deferred-acceptance algorithm, with courses keeping the current set of highest bids and rejecting those that don't fit. This algorithm stops when no student is rejected, at which point each student can take the courses which hold her name.

Krishna & Ünver (2008) showed that based on the theory and the results of their field experiments, asking for a rank ordered list as well as a bid list increases the efficiency of results of the GS algorithm in comparison with BPM. However, getting two sets of information which may not agree each other may in reality increase the risk of misreporting preferences and encourages students to behave strategically. Furthermore, the GS method like other bidding mechanisms at times results in solutions assigning no course to some students.

A fifth algorithm for CAP was introduced by Budish (2011), the Approximate Competitive Equilibrium for Equal Income (A-CEEI) Mechanism. In A-CEEI, students report their ordinal preferences over all possible schedules of courses and they have the option to report either their additive or non-additive complementary/substitutability preferences, depending on the specific implementation. Then, it assigns a random approximately-equal budget to students and allocates courses through a series of optimizations based on student preferences, student budgets, and course prices, emulating a Walrasian-style price equilibrium. Though several algorithmic variations are possible, one concrete implementation uses a metaheuristic Tabu search method to find a nearly-optimal price set, which approximately equates the number of available seats of each course (supply) with its demand. The course demand for each price set is calculated by allocating each course seat to a student who wants it and can afford it. The final allocation of courses is achieved when the error, the difference between allocated and available seats, is less than a predetermined tolerance. Accordingly, A-CEEI allows infeasible solutions in which the number of assigned students may exceed the course capacity, a problem to be sorted out after the mechanism. Also, similar to the draft mechanism, this method randomly prioritizes students, with randomly generated budget inequality directly affecting the results, which should in principle be avoided where possible.

Furthermore, finding an A-CEEI price-set that approximately balances the demand and supply of the market for real-size problems can be a very time consuming computational process. Indeed, because we were unable to create a fast implementation of A-CEEI approach subject to non-overlap constraints, we were unable to conduct a direct comparison of A-CEEI to our own mechanisms, which in contrast were easily implemented hundreds of times in our simulations. Further discussion of A-CEEI relative to the current context is provided in appendix C.

Along with these main algorithms, recently two other streams of research have been presented to solve CAP. In ***the*** first stream, another definition of CAP has been studied. Diebold et al. (2014) defined CAP differently and assumed that each course or organizer also has preferences over students. Based on their definition, CAP is a two-sided matching problem (Roth & Sotomayor, 1990). Therefore, they defined a stable matching of courses and students and compared the first come first served procedure and the GS mechanism with their mechanism. However, they did not consider overlapping constraints in their model and solutions. Nogareda & Camacho (2016) also had the same approach and defined CAP as a two-sided matching problem and they did not consider overlapping courses either.

In ***the*** second stream, first Budish et al. (2013) proposed a random allocation mechanism to solve the general matching problems including CAP. However, their mechanism cannot consider the section and time-slot overlapping constraints simultaneously. Also, design of their mechanism for multi-unit market is similar to Budish (2011) and both need to find a set of prices balancing supply and demand and clear the market. Finding this set of prices as we discussed for Budish (2011) is computationally very difficult. Recently, Akbarpour & Nikzad (2015) and Nguyen et al. (2015) proposed other random allocation mechanisms which similar to Budish (2011) and Budish et al. (2013) terminate with infeasible solutions. Here, our new methods all compute feasible solutions.

3. New algorithms and metrics for CAP

The course allocation problem (CAP) consists of a set of courses, some of which may have more than one section/time offered. Define the set $C = \{1, .., m\}$ as set of all course-sections offered in the market in which each course-section j has q_j seats. The course-section seats are allocated to a set of students $I = \{1, .., n\}$, each of whom can take at most k courses in each semester.

By standard assumptions of rationality, each student i has real cardinal preferences (utility) over course-sections u_{ij}. In any of our cardinal-eliciting mechanisms, she submits her cardinal preferences (i.e., bids) $b_{ij} \geq 0$ for each course-section with $\sum_j b_{ij} \leq 1000$. (As is standard in utility theory, affine transformations do not affect preferences, so it is always possible to normalize an approximately truthful revelation to this 1000 point bidding scale. Higher budgets would allow higher precision.)

In some mechanisms ordinal preferences are submitted; each student i submits her integer ordinal preferences over course-sections r_{ij} , where higher r_{ij} shows a more preferred course-section. (Though lower-is-better seems more common in plain language, with $r_{ij} = 1$ indicating the first-best course, etc., this approach is not practical for optimization, since non-allocation would seem lower, and therefore better.) For our computations, we let $0 \leq r_{ij} \leq 100$, though other upper bounds provide similar results. Unlike the GS algorithm for CAP in Sönmez and Ünver (2010) we require consistency among any student's course ranks and bids, i.e., $\forall j, j' , r_{ij} \leq r_{ij'} \Leftrightarrow b_{ij} \leq b_{ij'}$. In practice, for the new algorithms which use both cardinal and ordinal measures, a student need only submit her cardinal preferences, and then her ordinal preferences may be inferred due to this consistency requirement.

Each student will get a feasible bundle of course-sections S_i in which no two course-sections overlap, none are different sections of the same course, and the number of course-sections in the bundle is less than or equal to k. We define the binary decision variable x_{ij} to 1 if and only if a seat in course-section j is assigned to student i.

To define forbidden course overlap, let binary parameter $O_{jj'}$ equal 1 if and only if the course-section j and j' are overlapping course-sections in time or they are different sections of a same course. Therefore, a feasible bundle of course-sections, S_i , is defined by the corresponding x_{ij} , which must satisfy $\sum_{j=1}^{m} x_{ij} \leq k$ and $x_{ij} + x_{ij'} \leq 1$ for all (j, j') with $O_{jj'} = 1$. Although it is possible to require transitivity, in which $(O_{jj'} = 1)$ and $(O_{j'j''} = 1)$ would imply $(O_{jj''} = 1)$, we do not make such an assumption.

(Computational experiments that included a transitivity assumption were easier to solve, but we abandoned this line of enquiry, given that overlap based on the preponderance of non-transitivity in practice; one course might overlap with another in time, but not an alternative section of the same course, for example.)

The utility of the set S_i is defined as the sum of the utility of its members, i.e., we assume additive utility over course sections. Accordingly, the "best bundle of courses" for student i or S_i^* is also defined as a feasible bundle of course-sections for student i which maximizes her total utility.

For round-by-round algorithms, the letter t will be added as a round index subscript, giving us q_{jt} for the available seats in course j in round t and x_{ijt} for an allocation of a seat in course j to student i in round t. The set of eligible course-sections for student i in round t is denoted E_{it}. For round 1, we initialize $E_{i1} = \{j: b_{ij} > 0\}$, with $E_{i,t+1}$ formed by successively removing courses no longer in play at the end of each round. Also, p_{jt} is defined as the price of course j in round $\boldsymbol{t}$, which is needed for one of the proposed algorithms. For simplicity, in the rest of this paper the term "course" refers to the "course-section" unless it is explicitly mentioned.

Turning our attention from notation to the specific metrics of interest, market efficiency is measured by the following integer programming formulation:

$$Max \sum_{i,j} u_{ij}\, x_{ij} \tag{1}$$

$$\sum_i x_{ij} \le q_j \quad \forall j \tag{2}$$

$$\sum_j x_{ij} \le k \quad \forall i \tag{3}$$

$$x_{ij} + x_{ij'} \le 1 \quad \forall j, j' \text{ with } O_{jj'} = 1 \tag{4}$$

$$x_{ij} \in \{0,1\} \quad \forall i, j \tag{5}$$

The objective is to maximize total (cardinal) utility for the market, with no more seats awarded than are available in any course (2), no more than k courses per student (3), and no student taking conflicting courses (4). Any allocation mechanism satisfying (2) -(5) can be compared using the following natural efficiency and fairness metrics as discussed in §1 and summarized in Table 1. Typically, in practice, market efficiency based on true utility u_{ij} is not possible, since only bid values b_{ij} are given. So for any of these metrics, both revealed and actual metrics are possible based on the situation, the latter typically only possible via simulation or controlled experiment.

	Efficiency	Fairness	
Category	Larger ⇒ better	Smaller ⇒ better	
Cardinal	$\sum_{i,j} u_{ij}\, x_{ij}$	$\mathrm{Range}_i(\sum_j u_{ij}\, x_{ij})$	$\mathrm{StDev}_i(\sum_j u_{ij}\, x_{ij})$
Ordinal	$\sum_{i,j} r_{ij}\, x_{ij}$	$\mathrm{Range}_i(\sum_j r_{ij}\, x_{ij})$	$\mathrm{StDev}_i(\sum_j r_{ij}\, x_{ij})$
Binary	$\sum_{i,j} x_{ij}$	$\mathrm{Range}_i(\sum_j x_{ij})$	$\mathrm{StDev}_i(\sum_j x_{ij})$

Table 1 – Efficiency and Fairness Metrics

3.1. Multi-Round Top-Trading-Cycle Algorithm (TTC) for CAP

The top-trading-cycle algorithm was developed to solve the house-allocation problem, which finds a new assignment of houses among a set of house-owners who are selling their houses and each owner has preferences over other houses in the market. In this algorithm, each owner points to her most preferred house, and each house points to the highest offer. If there is a cycle (i.e., a house points to an owner and the owner also points back to the same house) the owner is assigned to the house, and both are removed from the market. The process then reiterates with the remaining participants. Another version of this algorithm assigns students to schools and has been used in practice. In the current paper, we modify this algorithm to solve the course allocation problem.

Our new top-trading-cycle algorithm (TTC) for CAP attempts to overcome the weaknesses of BPM and the draft mechanism while keeping their strengths. BPM considers the cardinal bids of students and avoids random favoritism of students when it is not needed; however, it may assign all top preferred courses to some students and assign no course to others. On the other hand, the draft mechanism randomly prioritizes students and assigns courses to students based on the random order. Nevertheless, it considers fairness as no student can get her $t + 1^{st}$ course before all other students get their t^{th} course. TTC is developed to combine these strong points. Accordingly, it assigns one course in each round to each student, but the winning students are defined based on the students' bid values and not based on random numbers.

The details of this algorithm are as follows:

1- Each student is given a fixed budget (e.g., 1000 points) and submits bid points b_{ij} for each course with the sum not exceeding the budget.

2- Round t, the algorithm acts on behalf of each student as follows:

   I. Each student points to $j_i^* = argmax_{j \in E_{it}} b_{ij}$ and offers the bid amount ($b_{ij_i^*}$) for it.

   II. Each course accepts up to q_{jt} of the highest offers and rejects any remaining offers. Then capacities are updated: $q_{jt+1} = q_{jt} - \sum_{i=1}^{N} x_{ijt}$.

   III. $E_{i,t+1}$is calculated based on the following recursive equation:

$$E_{i,t+1} = E_{it} - \{j_i^*\} - \{j: x_{ij_i^*t} = 1 , O_{jj_i^*} = 1\} - \{j: q_{jt+1} = 0\}$$

   IV. If there are rejected students, they will repeat steps I to III until no more student is rejected. Accordingly, all students take <u>one</u> more course in round t or leave the market because there are no courses in their E_{it}.

3- Step 2 repeats until all students take k courses or leave the market.

Example 1. Consider a set of students $I = \{S1, S2, S3, S4\}$, and courses $C = \{C1, C2, C3, C4, C5\}$ with $\{2,3,3,2,2\}$ seats, respectively, where C1 and C4 are overlapping courses and each student can take at most three courses. Table 2 shows the submitted student bid lists (cardinal preferences) and ranked-ordered lists (ordinal preferences) for the offered courses. In Example 1, we assume that $1 \leq r_{ij} \leq 5$ and higher r_{ij}represents more preferred courses.

Under TTC, students first point to their top ranked course. Therefore, C1 receives two offers from S1 and S4 and based on its capacity accepts theses offers. At the same time, C3 and C4 receive one offer from S2 and S3, respectively, and they accept these offers. Thus round 1 ends, with one course assigned to each student and C1 with no more seats leaving the market. The updated course capacity list is {0,3,2,1,2}.

Student 1		Student 2		Student 3		Student 4	
Rank	Bid List	Rank	Bid List	Rank	Bid List	Rank	Bid List
C1	400	C3	256	C4	245	C1	251
C3	230	C2	252	C1	243	C3	242
C4	200	C4	246	C3	240	C2	235
C2	150	C1	245	C2	230	C4	201
C5	20	C5	1	C5	42	C5	71

Table 2 - Student rankings and bid lists for Example 1

In round 2, C2 accepts one offer from S2, but C3 receives three offers from S1, S3, and S4 which is more than its capacity. Then C3 fills its two available seats from the two highest offers (242 and 240) and rejects the lowest one (200). Therefore, S1's offer for C3 is rejected, but she gets another chance and her next course in her rank ordered list. She points to the fourth course in her rank ordered list, skipping C4 due to overlap with C1 already in her schedule, and her offer for C2 is accepted. Each student has two courses so round 2 ends, with the updated course capacity list {0,1,0,1,2}.

In round 3, S1 and S2 point to their next highest ranked course among available course, C5 and C4, respectively, and since there are enough seats, their offers are accepted. At the same time, C2 receives two offers from S3 and S4, but with only one available seat, it only accepts the higher bid offer (235) and rejects the lower one (230). Accordingly, S3 who is rejected by C2 gets another chance and points to his next highest ranked course with available seats, C5, and since it has an available seat, he takes the course. With no course seats available, the final course assignment for Example 1 and the total rank of assigned courses to each student is shown in Table 3. Also, Table 4 shows the efficiency and fairness metrics for Example 1 solved by TTC.

Assigned Courses	Student 1	Student 2	Student 3	Student 4
Round 1	C1	C3	C4	C1
Round 2	C2	C2	C3	C3
Round 3	C5	C4	C5	C2
Total Cardinal Utility	570	754	527	728

Total Ordinal Utility	8	12	9	12
Total Binary Utility	3	3	3	3

Table 3 – Final solution to Example 1 solved by TTC

	Efficiency	Fairness	
		Range	St. Dev.
Cardinal	2579	227	97.88
Ordinal	41	4	1.79
Binary	12	0	0

Table 4 – Efficiency and Fairness Metrics of Final solution to Example 1 solved by TTC

3.2. Second-Price Algorithm (SP) for CAP

As mentioned above, the VCG mechanism can elicit truthful bids from bidders in an auction, where monetary bids are used to compute winners and payments for goods received. However, in CAP, monetary transfers cannot be used, making it more difficult to control incentives; in bidding with points (a fictional currency used only for the CAP process) receiving low prices or keeping unused points has no value to a bidding student, and therefore incentives are drastically different.

But in a multi-round setup, we explore the possibility that returned points may retain some currency value from round to round, encouraging truthful bidding. For example, suppose a bidder would like to bid aggressively for her top course, say 350 points, only to find later that she could have bid 250 and still would have been awarded the course. Retrospectively, she would have wished to have bid lower for this course so that more of the bid points could have been used on subsequent course bids. This suggests an incentive for her to strategically lower her truly high value for the course. But if a round-by-round mechanism returned points that exceeded the threshold necessary to "buy" the course with points (i.e. gave her back the 100 points excess), she would not have such an incentive.

Here, we explore this notion of returning points, similar to discounts in a second-price auction, to be used across rounds. In a setting with unit-demand bidders, the VCG mechanism becomes a uniform-price auction in which the top q bidders win the q available items and pay the amount of the $q+1^{\text{st}}$ bid, which we emulate in each round of our multi-round set-up. The resulting algorithm is called the second-price algorithm for CAP, or SP for short. The algorithm is very similar to TTC, but with step 2 part II replaced as follows with $\mathrm{II}^{\mathrm{SP}}$:

$\mathrm{II}^{\mathrm{SP}}$. Each course accepts up to q_{jt} of the highest offers and rejects any remaining offers. Then capacities are updated: $q_{jt+1} = q_{jt} - \sum_{i=1}^{N} x_{ijt}$.

If $q_{jt+1} = 0$, then let price p_{jt} equal the highest rejected offer for j, else let $p_{jt} = 0$. Return unspent points to use on the next most preferred course, i.e., let $j_i^{next} = \ argmax_{j\in E_{it}-\{j_i^*\}} b_{ij}$ and let $b_{ij_i^{next}} = b_{ij_i^{next}} + b_{ij_i^*} - p_{jt}$.

We next consider the same market as discussed in Example 1, but now under SP. As before, C1 receives two offers, C3 receives one offer, and C4 receives one offer in the first round. Since there are no rejected offers, the prices of all courses in this round are set to zero and the points of all offers are added to the next course in their rank ordered list. In round 2, courses begin to reach capacities and thus charge prices for courses, only returning the unused portion of each bid, which is then used for the next set of offers. These next offers will occur before round 2 ends, because some students were rejected from their second favorite course and must be awarded a course before the next round begins.

We omit the full details, but the results of SP on Example 1 with efficiency and fairness metrics are given in Tables 5 and 6, respectively. Note that relative to the TTC results in Table 3, this final allocation awards the same number of courses and has the same ordinal efficiency (same total rank). However, other measures tell different stories: the ordinal fairness metrics (both range and standard deviation of rank) are lower (better) for SP, while the cardinal fairness metrics (range and standard deviation of cardinal utility) are both lower (better) for TTC. Also, total cardinal utility is larger for SP, indicating overall that it is difficult to rank allocations/algorithms based on the many design metrics, even for a very small (oversimplified) problem.

Assigned Courses	Student 1	Student 2	Student 3	Student 4

Round 1	C1	C3	C4	C1
Round 2	C3	C2	C3	C2
Round 3	C2	C4	C5	C5
Total Cardinal Utility	780	754	527	557
Total Ordinal Utility	11	12	9	9
Total Binary Utility	3	3	3	3

Table 5 – Final solution to Example 1 solved by SP

	Efficiency	Fairness	
		Range	St. Dev.
Cardinal	2618	253	113.37
Ordinal	41	3	1.30
Binary	12	0	0

Table 6 – Efficiency and Fairness Metrics of Final solution to Example 1 solved by SP

3.3. The optimized multi-round algorithms: TTC-O and SP-O

The TTC and SP algorithms locally maximize the benefit of each student in each round in a greedy fashion, not considering that a better outcome may be available for the whole market, even within the same round. Instead, at times exchanging a student's r^{th} ranked course with her $(r+1)^{st}$ course may benefit two or more students. In this case, the total rank for the whole market can be improved, with two or more made better off and one student worse off. Considering the benefit to the entire market in each round is made possible in the following optimization versions of the TTC and the SP mechanisms, which we call TTC-O and SP-O.

In both methods (and as in our fifth and final method, OC) we make use of a two-stage optimization, first finding the maximum possible total rank, and then, maximizing total bid points among all rank-optimal solutions, since there are typically many. The optimizations for TTC-O and SP-O occur in each round, in contrast to OC described below.

Multi-Round Top Trading Cycle Algorithm, Optimization Version (TTC-O)

In round t we solve optimization model (6) -(9) followed by (10)-(11):

$$Z_{t1} = Max \sum_{j,i \in E_{it}} r_{ij}\, x_{ijt} \qquad (6)$$

$$\sum_i x_{ijt} \le q_{jt} \quad \forall j \qquad (7)$$

$$\sum_{j \in E_{it}} x_{ijt} \leq 1 \quad \forall i \tag{8}$$

$$x_{ijt} \in \{0,1\}, x_{ijt} \in E_{it} \tag{9}$$

$$Z_{t2} = Max \sum_{j,i \in E_{it}} b_{ij}\, x_{ijt} \tag{10}$$

(7)- (9)

$$\sum_{j,i \in E_{it}} r_{ij}\, x_{ijt} = Z_{t1} \tag{11}$$

Notice that (6) -(9) represent an instance of the transportation problem, known to have a totally-unimodular constraint matrix, thus making it easy to solve as a relaxed integer program. The addition of constraint (11) to lock in the total rank at its maximum value and the changing the objective to (10) in the secondary optimization do not make the problem harder in practice.

Consider again the market discussed in Example 1. In the first round of TTC-O, just as in TTC, courses C1, C3, C4, C1 are assigned to S1, S2, S3, S4, respectively. (Since all can get their top-ranked course, it is optimal to do so in the first round.) But, in round 2, the TTC-O method differs, with C3, C2, C3, C2 assigned to S1, S2, S3, S4, respectively. In TTC, S4's bid for C3 beats out S1's offer for this course, but TTC-O recognizes that denying the seat to S4 results in her getting a course of only one lower rank, whereas denying the seat to S1 would push her two steps down her rank list, a worse move overall for total rank in this round. (The third round also proceeds differently, though details are omitted.) The final assignment of courses to students with the TTC-O algorithm and efficiency and fairness metrics are presented in Table 7 and 8, respectively.

Comparing Table 7 to the results of TTC and SP (Tables 3 and 5), we see a third distinct allocation of courses to students delivering the same ordinal efficiency of 41. The results in Table 7 have a higher (better) cardinal utility, also we see that its allocation performs at least as well if not better in every other fairness and efficiency metric.

Assigned Courses	Student 1	Student 2	Student 3	Student 4
Round 1	C1	C3	C4	C1
Round 2	C3	C2	C3	C2
Round 3	C5	C4	C2	C5

Total Cardinal Utility	650	754	715	557
Total Ordinal Utility	10	12	10	9
Total Binary Utility	3	3	3	3

Table 7 – Final solution to Example 1 solved by TTC-O or SP-O

	Efficiency	Fairness	
		Range	St. Dev.
Cardinal	2676	197	74.58
Ordinal	41	3	1.09
Binary	12	0	0

Table 8 – Efficiency and Fairness Metrics of Final solution to Example 1 solved by TTC-O or SP-O

A Second-Price Algorithm for CAP, Optimization Version (SP-O)

In the SP-O algorithm, at the end of each round the excess points are returned to students and applied to their next round offers. To calculate the extra points, we need to determine the price of each course in each round. The dual optimization problem (12)-(14) followed by (15)-(16) at the end of round t helps us to find the course prices in this round.

$$W_{t1} = Min\ (Z_{t1}\ D + \sum_j q_{jt}\, p_{jt} + \sum_i v_{it}) \tag{12}$$

$$p_{jt} + v_{it} + r_{ij}\, D \geq b_{ij} \tag{13}$$

$$p_{jt}, v_{it} \geq 0\ , D\ is\ unrestricted \tag{14}$$

$$W_{t2} = Min\ \sum_j q_{jt}\, p_{jt} \tag{15}$$

(13)- (14)

$$Z_{t1}\ D + \sum_j q_{jt}\, p_{jt} + \sum_i v_{it} = W_{t1} \tag{16}$$

In these models, p_{jt} is the price of course j in round t and v_{it} can be described as the value of student i in round t in the market. Also, in (15) we find a set of prices which minimize the weighted summation of course prices. It means that we are interested to decrease the price of courses with fewer seats less than courses with more seats. This objective function along with (16) which locks (12) at its optimum value result in finding unique optimum values for the course prices in each round.

Consider again the market discussed in Example 1. Since this is a very small example, returning points to students does not change the final solutions and the assigned courses in each round of TTC-O and SP-O and their final solutions are the same. Therefore, the final assignment of courses to students with SP-O algorithm and efficiency and fairness metrics are presented in Table 7 and 8, respectively.

3.4. Ordinal-then-Cardinal (OC) Algorithm for CAP

In the Ordinal-then-Cardinal (OC) algorithm for CAP, we maximize ordinal utility followed by maximizing cardinal utility among rank-maximal solutions as in TTC-O and SP-O, but here performing this two-part optimization once for the whole market instead of in each round. To define these models formally, we use the general market feasibility constraints (2)-(5).

$$Z_1 = Max \sum_{i,j} r_{ij}\, x_{ij} \tag{17}$$

(2)-(5)

$$Z_2 = Max \sum_{i,j} b_{ij}\, x_{ij} \tag{18}$$

$$\sum_{i,j} r_{ij}\, x_{ij} = Z_1 \tag{19}$$

(2)- (5)

Tables 9 and 10 display the final assignment of courses to students as well as efficiency and fairness metrics under OC using a maximum schedule size of k = 3, respectively. However, if k changes and increases to 4 the solution of OC will change while the solution of other presented algorithms stay the same. Tables 11 and 12 show the final assignment as well as the efficiency and fairness metrics of courses when k = 4, respectively. Both solutions achieve a total rank of 42, one unit higher than any previously discussed allocation, with the former showing that it is possible to achieve this better solution while maintaining equal schedule sizes (binary fairness metrics equal to zero). While this small ***change*** shows a potential issue of fairness in OC results, our simulation results show that OC algorithm is able to achieve the lowest binary fairness metric among all algorithms.

	Student 1	Student 2	Student 3	Student 4
Assigned Courses	C1, C3, C5	C3, C2, C4	C4, C2, C5	C1, C3, C2

Total Cardinal Utility	650	754	517	728
Total Ordinal Utility	10	12	8	12
Total Binary Utility	3	3	3	3

Table 9 – Final solution to Example 1 solved by OC using $k = 3$.

	Efficiency	Fairness	
		Range	St. Dev.
Cardinal	2649	237	92.18
Ordinal	42	4	1.66
Binary	12	0	0

Table 10 – Efficiency and Fairness Metrics of Final solution to Example 1 solved by OC using $k = 3$

	Student 1	Student 2	Student 3	Student 4
Assigned Courses	C1, C3	C3,C2,C4	C4,C2,C5	C1, C3, C2,C5
Total Cardinal Utility	630	754	517	799
Total Ordinal Utility	9	12	8	13
Total Binary Utility	2	3	3	4

Table 11 – Final solution to Example 1 solved by OC using $k = 4$.

	Efficiency	Fairness	
		Range	St. Dev.
Cardinal	2700	282	110.23
Ordinal	42	5	2.06
Binary	12	2	0.71

Table 12 – Efficiency and Fairness Metrics of Final solution to Example 1 solved by OC using $k = 4$

3.5. Pareto (in)efficiency

Given these 5 new algorithms, it is natural to ask whether each satisfies Pareto efficiency with respect to submitted preferences. It is easy to see that optimization of over the entire market as in (1)-(5) results in Pareto efficiency with respect to the objective function used, cardinal, ordinal, or binary. (This follows because the existence of a feasible Pareto improvement would contradict optimality.) Thus, OC guarantees ordinal Pareto efficiency due to its first optimization. Though an OC solution may not be cardinal Pareto efficient, any Pareto improving solution in cardinal utility must necessarily degrade the total ordinal utility of the market (harming at least one student in ordinal preferences) by virtue of its ordinal then cardinal optimization. Thus solutions that improve the situation on both metrics are not available. Sadly, the same guarantees cannot be made for the other mechanisms, as can be shown via counter Example 2 described in Table 13.

Example 2. Consider a market with two students S1 and S2 each of whom can take at most 2 courses (k = 2) and a set of five courses $C = \{C1, C2, \dots, C5\}$, each with one seat. Let C1 and C3 be overlapping courses, and let the preferences of S1 and S2 be as displayed in Table 13.

	Bid lists	
Courses	Student 1	Student 2
C1	385	380
C2	320	350
C3	180	100
C4	105	120
C5	10	50

Table 13 – Bids showing Pareto inefficiency under course overlap

Under each of the mechanisms discussed here, with the exception of OC, a Pareto inefficient solution is selected, considering either ordinal or cardinal utility. Under the draft mechanism, BPM, TTC, or SP, student 1 gets {C1, C5} while student 2 gets {C2, C4}. Alternatively, under TTC-O and SP-O, students 1 and 2 {C1, C4} and {C2, C3}, respectively. But OC, on the other hand awards them {C2, C3} and {C1, C4}, respectively, which is a (cardinal or ordinal) Pareto improvement over either outcome. Note also that straightforward cardinal-utility maximization results in an alternate cardinal Pareto efficient solution, {C3, C4} and {C1, C2}, but OC tends to reject this more lopsided solution, which seems to place too much weight on small cardinal differences as compared to the relative equity implied in ordinal comparisons. Note also that this latter cardinal Pareto efficient solution fails to be ordinal Pareto efficient, while the OC solution is Pareto efficient in both senses.

4. Efficiency and Fairness Results

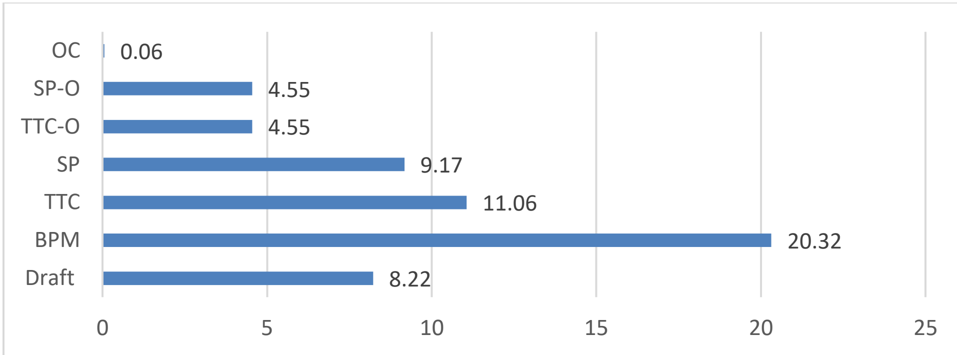
In order to assess the proposed algorithms and compare their results with BPM and the draft mechanism, we randomly generated 100 sample markets. Each consists of n = 900 students, each of whom can take up to k = 6courses and 83 courses some with multiple sections for a total of m = 112 course-sections. (We assumed 2 courses having five sections, one course with four sections, two courses with three sections, fourteen courses with two sections, and 64 courses with one section.) These features were selected to roughly approximate the characteristics of offered courses at Harvard Business School, a prominent user of algorithmic course-allocation. Accordingly, the capacity of 10%, 30%, and 60% of course-sections were drawn from discrete uniform distributions with probabilities of {15, 25, 35}, {40, 50, 60}, and {70, 80, 90, 100}, respectively. To establish overlap among course-sections, 8 weekly time slots were generated, and course-sections were considered overlapping if they shared a time slot or were sections of the same course. In addition, similar to HBS, we generated 10 different major fields of study to group correlated courses, and based student preferences on these fields of study. Further details on how these preferences were generated are given in Appendix A. All experiments in this paper involving optimization used CPLEX version 12.1 on a 2.53GHz machine.

In this first set of experiments, we assumed that the generated cardinal and ordinal preferences are true student preferences and they have not been manipulated to improve one's allocation. (In §5, we will relax this truthfulness assumption in a second set of experiments.) Here, we present and interpret values for the efficiency and fairness metrics proposed in Table 1.

4.1. Binary Efficiency and Fairness

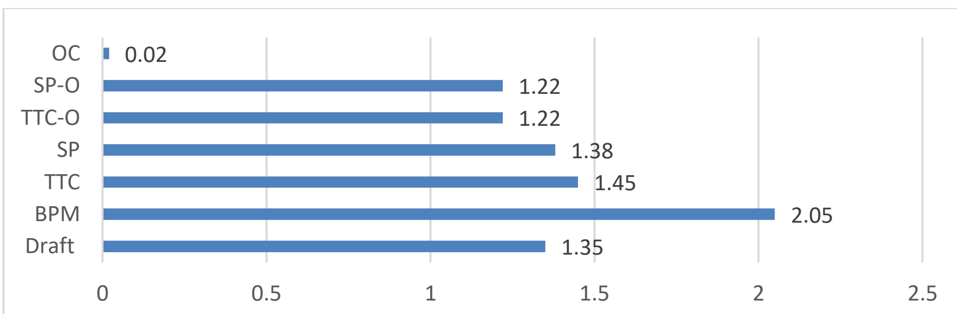
Recall that our notion of binary utility measured a 1 for an assigned seat a 0 otherwise. Averaging total binary utility of the market over each of 100 simulated markets therefore shows us the average number of assigned seats by each mechanism. With 900 students seeking 6 courses in each case, each market has 5400 opportunities to assign students to a course. For easier interpretation, Figure 1 shows the average number of missed opportunities (amount below 5400) for each algorithm. Clearly, the optimization-based algorithms do a better job of getting students into seats, with OC doing a very good job, missing only 6 opportunities out of 540,000 opportunities to place students.

Figure 1 - The average number of missed assignment opportunities per algorithm



Turning next to the range of binary utility across students as a metric of fairness, we note that in almost every random trial, at least one student received a full six-course schedule. Thus the range is most indicative of how the worst student is treated in terms of the number of courses she is assigned. Figure 2 shows the average of the inter-student range across simulation instances.

Figure 2-The average range of binary utility per student per algorithm

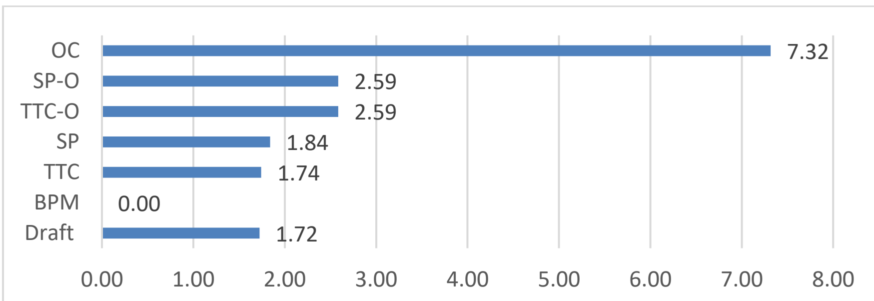


As Figure 2 depicts, in BPM the range is 2.05 on average, i.e., the student with lightest schedule gets only 3.95 courses on average. While BPM assigns 3 or fewer courses to at least one student for 12 out of 100 instances, this occurred for only one instance under TTC or SP. But, again under this metric, the optimization-based approaches do best; none ever assigned fewer than 4 courses at worst, and OC never assigned fewer than 5 courses to students in any instance. Indeed, in 98 of 100 instances, OC assigned 6 courses to every single student, while BPM did not do so (complete 5400 assignments) in even one instance. Results for the standard deviation of binary utility were qualitatively quite similar to those presented in Figure 2, and are thus omitted.

4.2. Ordinal Efficiency and Fairness

Figure 3 shows the simulation results of the different mechanisms based on the average total rank (ordinal efficiency) of assigned courses. As BPM scored the worst on this metric, the figure shows amounts above this worst benchmark case. Clearly again, the optimization-based methods dominate the other algorithms, with OC in a strong lead, getting much better outcomes in terms of how students order the courses.

Figure 3 - The average ordinal utility per student per algorithm, showing amounts above BPM benchmark with average ordinal utility of 548.15 per student.



Equity in the rank of assigned courses gives us our next measure of fairness across students, as shown in Figure 4. Since in almost all instances the maximum total rank of assigned courses to a student is equal to the maximum possible 585 for all algorithms considered, the range of ordinal utility is indicative of the total rank of the worst student and how her total rank compares to this maximum. Note that the total rank equal to 585 means that a student could take her (100, 99, 98, 97, 96, 95) ranked courses. In other words, she could take the 6 highest ranked courses on her ordered list. Moreover, the worst total rank possible for a student with a full schedule in these simulations (given the artificial 100-point ordinal basis and bidding on a maximum of 35 different courses) is 411, corresponding to her ordinal objective value when she receives her 30^{th}, 31^{st}, … 35^{th} best courses. Thus, ranges greater than $585 - 411 = 174$ represent a worst student necessarily getting a reduced course load. Also note that this measure does grow quickly: a worst student getting her 2^{nd}, 3^{rd}, … 7^{th} courses (her second best possible schedule, the second best possible market measure) indicates a range of 6, while simply removing one course from the worst-off student increases the range by at least 66 (the ordinal objective value of a 35th best course.)

Figure 4 - The average range of ordinal utility per student per algorithm

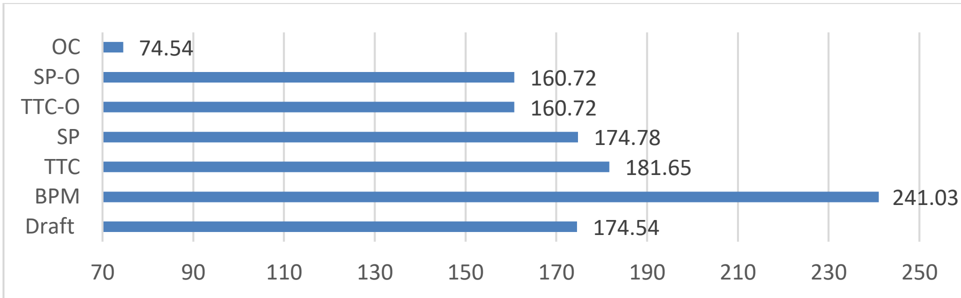
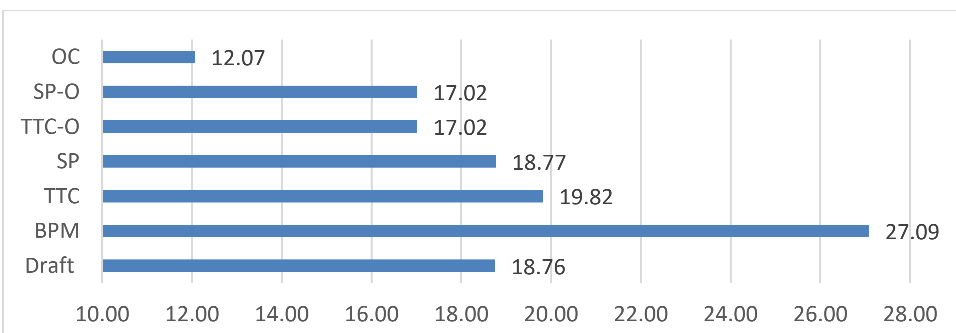


Figure 4 again shows the same general pattern, with BPM scoring worst, optimization-based algorithms doing best, and OC in a clear lead compared to all others. Looking at Figure 5, which shows average ordinal fairness performance via the standard deviation, we see a similar pattern, but with less contrast on average across algorithms than the extreme cases measured by the range.

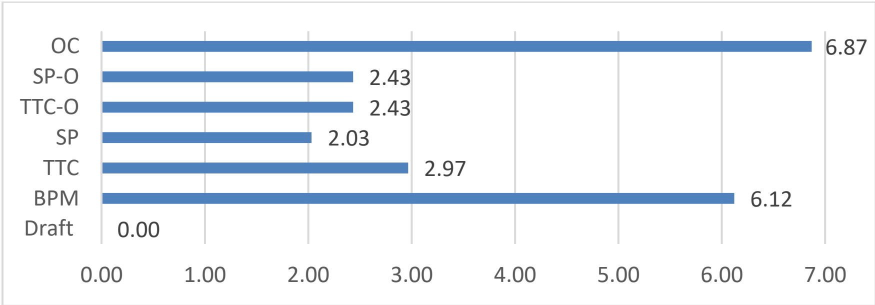
Figure 5 – Average standard deviation of ordinal utility per student per algorithm



4.3. Cardinal Efficiency and Fairness

Figure 6 displays the average cardinal utility from each student's assigned courses under each mechanism. As Figure 6 depicts, the draft mechanism achieves the lowest average cardinal utility, with OC and BPM achieving the first and second best on this metric, respectively. Though BPM does well on this metric, comparing Figure 6 to previous metrics, we see that BPM gets a high cardinal utility by grabbing at it greedily, sacrificing fairness and efficiency among the previous metrics. In this context, OC proves that this sacrifice is unjustified; it is possible (as OC does) to score better than BPM in all categories.

Figure 6 – The average cardinal utility per algorithm, amounts above Draft benchmark with average cardinal utility 295.93



Indeed, Figure 7 displays the average range of cardinal utility, where we see that BPM gets a high overall cardinal utility by increasing the inequality between the top and the bottom of the market. Again we see optimization-based algorithms dominating the others, but here SP-O and TTC-O edge out OC, in contrast to the previous results we have seen. OC seems willing to increase the range of cardinal preferences in order to keep the ordinal measures more equitable, consistent with its ordinal-then-cardinal formulation.

Figure 7 – The average range of cardinal utility per algorithm

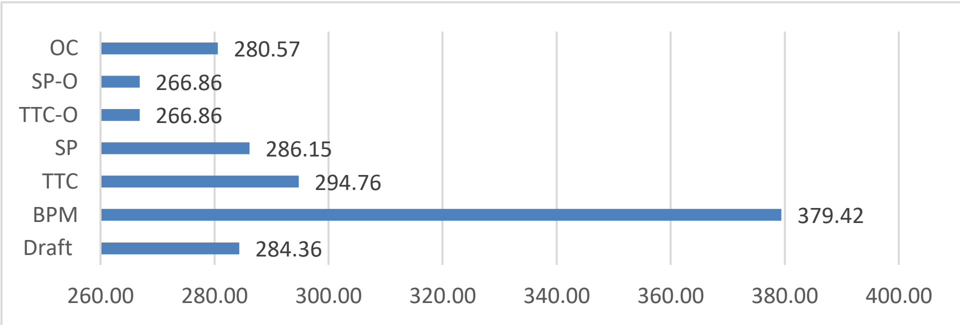
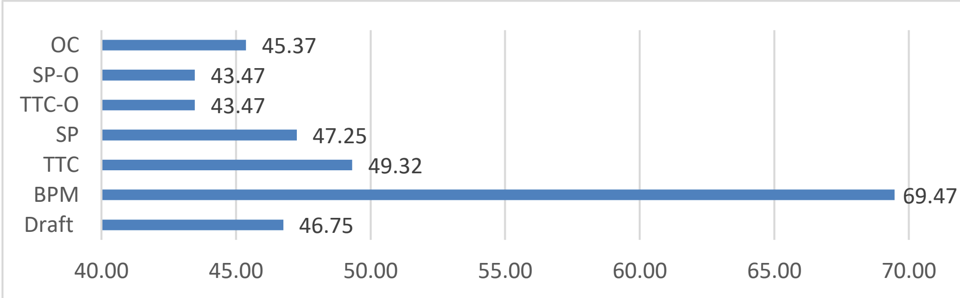


Figure 8 again shows the standard deviation painting a similar picture to that of the range, again favoring the optimization-based algorithms and relaying bad news for BPM.

Figure 8 – Average standard deviation of cardinal utility per algorithm



5. Incentive Compatibility Results

To measure the effect of strategic behavior on the results of the various mechanisms discussed so far, we adapt and build on a process presented by Kominers et al. (2010) and used by Budish & Cantillon (2011). As in those papers, we model student preferences as arising from a combination of personal preferences (simulated randomly for each student, as in §4) and a common-value component (drawn once for each course in the market). Each student's true preferences are then modeled as the average of her personal signal for the value of the course and the market's common-value signal of the value of the course. This leads to a natural mechanism for simulating strategic behavior: a strategic student adjusts the relative weight between the personal signal and the common-value signal, placing greater weight on the common-value signal than on her true preferences.

This method for measuring incentives to deviate from truth-telling is an imperfect heuristic, but it gives a realistic basis for comparing various mechanisms in this highly complex environment, and is consistent with the notion of bounded rationality. Indeed, in real-life applications, it would be highly *unrealistic* to assume that students would have enough information about the preferences and bids of others in order to devise an optimal strategic bid. Even if a particular student had *complete and perfect knowledge* of the bids of others (which would be impossible in practice unless the bid submission system was somehow compromised) the problem of devising an optimal strategy under each of the mechanisms in question is typically a difficult and large NP-hard optimization problem.

If one instead tried to model beliefs over the preferences of others, the solution concept becomes the determination of a Bayes-Nash Equilibrium, which would be even more difficult to compute, and again out of the reach of a student in almost any realistic setting. (Though we considered these theoretical problems in our own notes, we leave a detailed study of this problem for future research.) So instead, we apply the benchmark paradigm explored in the literature, but here running the procedure for a broad set of parameter values, including several values for the percentage of students bidding strategically and for how much they manipulate their true preferences. This provides a robust investigation of the possibilities for student gain through deviation from truth-telling.

As in our first set of experiments, we find it beneficial to explore a number of simple metrics to interpret our results. These illustrate how variations in *the percentage of strategic students* and in *the intensity of their strategic behavior* change the outcomes for strategic and truthful students. In this section, we focus on ordinal utility and do not present the parallel results for binary and cardinal utility measures. We do this for brevity, and because cardinal results are similar but mildly harder to interpret, while binary utility does not shift drastically based on strategies (a strategic manipulation may shift *which* student gets a small and which a large bundle, without changing a binary metric, for example.)

To make our heuristic model of strategic behavior precise, consider equation (20), which is nearly identical to the approach presented in Kominers et al. (2010):

$$Bid_{ij} = (1 - w_i) * u_{ij} + w_i * CV_j \qquad (20)$$

In this equation, Bid_{ij} represents the *declared* cardinal utility of student i for course j, which is calculated as a simple weighted average of student i's personal utility component for course j (labeled u_{ij}) and the common value component of course j in the market (labeled CV_j). The personal parameter w_i controls the relative weighting of these two components. By assumption (as in the previous literature) a bidder who bids truthfully is defined as one who sets $w_i = 0.5$. When student i sets $w_i > 0.5$, she is effectively pretending to be more influenced by the common-value term than her true preferences dictate, causing her to bid more aggressively on popular courses (those with higher CV_j) and less aggressively on unpopular courses. Such a student is referred to as "strategic." Here, similar to our model in §4, students' cardinal and ordinal preferences are forced to be consistent, so that a single list of cardinal utility for each course j as in (20) is sufficient to generate a unique ordinal list for each student. As a robustness check, we also explored the settings where $\mathrm{w_i}$ of a truthful student was defined to be 0.25 and 0.75 and repeated all analysis. Our results showed that the conclusions remain largely unchanged for all algorithms. Incentive compatibility metrics of this analysis are discussed in Appendix B.

After an initial test of this common-value approach, we noticed systematic poor performance for students using the most straightforward implementation of the approach as described thus far. We then found improved payoff to strategic manipulation for students employing the non-truthful common-value adjustment *only* for course-sections in her "expanded best bundle," defined as her favorite feasible bundle together with any course-sections with overlap constraints incident to this bundle. Without this measure, many instances saw degraded performance when a high common-value course with low personal-value was focused on when an overall better course was available. The results presented here use this heuristically improved method of strategic manipulation.

Given this setup, we varied two parameters to explore the effect of strategic behavior. The first parameter, labeled P, indicates the percentage of students in the market that play strategically. (Hence, $(1-P)$ students bid truthfully, i.e., maintaining $w_i = 0.5$.) For each value of P, we also varied the w_i parameter for strategic students. With five values for $P = \{20\%, 40\%, 60\%, 80\%, 100\%\}$, and five values for $w_i = \{0.6, 0.7, 0.8, 0.9, 1.0\}$ of strategic students, each mechanism was tested under 25 different scenarios.

5.1. Effect of w_i when $P = 20\%$

To define the benefit of strategic behavior for each (P, w_i) scenario, we consistently compute the totally truthful benchmark scenario, in which $P = 0$. The "change in ordinal utility of each strategic student" is defined as the total (true) rank of assigned courses to her in the (P, w_i) scenario minus the total (true) rank in the totally truthful benchmark. The average of this metric over strategic students is shown in Figure 9 for varying w_i values under $P = 20\%$. The left graph in Figure 9 includes the results of all seven algorithms, but to scale better, the right one displays the results of all algorithms except BPM.

Figure 9 – Average change in ordinal utility of strategic students under $P = 20\%$

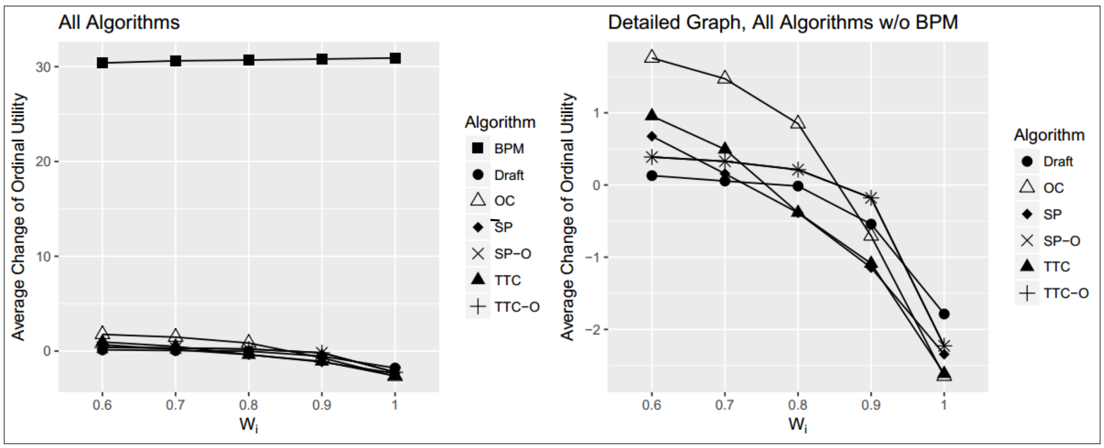


Figure 9 illustrates that as w_i increases, the average benefit of a strategic student decreases in all mechanisms except BPM. It also shows that among all algorithms, BPM is by far the most vulnerable to strategic manipulation, with the left graph showing relatively minor differences among all other mechanisms. Zooming in (the right graph), however, does show some subtle differences, with the draft mechanism showing the smallest potential benefit to deviation from truth-telling. All of the mechanisms on the right show a decreasing pattern, in which larger deviations (via larger w_i values) help the strategic students less, becoming a negative expected value (i.e., the deviation was worse than telling the truth) for a large enough $\mathrm{w_i}$. However, under BPM, strategic behavior never hurts students, indicating that this dishonest strategy is robust.

Overall though, to gauge the magnitude of these effects of strategic manipulation, it is important to remember that ordinal rank measures change rather quickly as the bundle of assigned courses changes. Recall from §4.2 that moving from one's best possible schedule to one's second best possible schedule results in a change of 6, for example, or that getting one more course (relative to truth-telling) raises ordinal utility by at least 66. Thus, seeing average changes under 2 ordinal units for each of the mechanisms on the right in Figure 9 indicate that these mechanisms are performing rather well in terms of providing very little benefit on average to misreporting preferences.

Next, we can measure the average change in ordinal utility between each (P, w_i) scenario and the $P = 0$ benchmark, but this time for students who remain truthful in the former case, indicating how much they are harmed by the other students' decision to play strategically

Figure 10 – Average change in ordinal utility of truthful students under $P = 20\%$

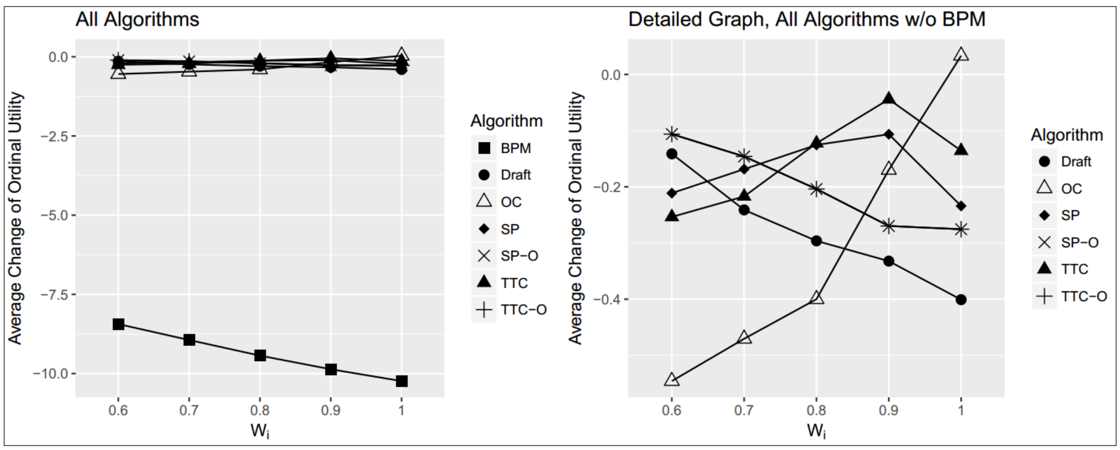
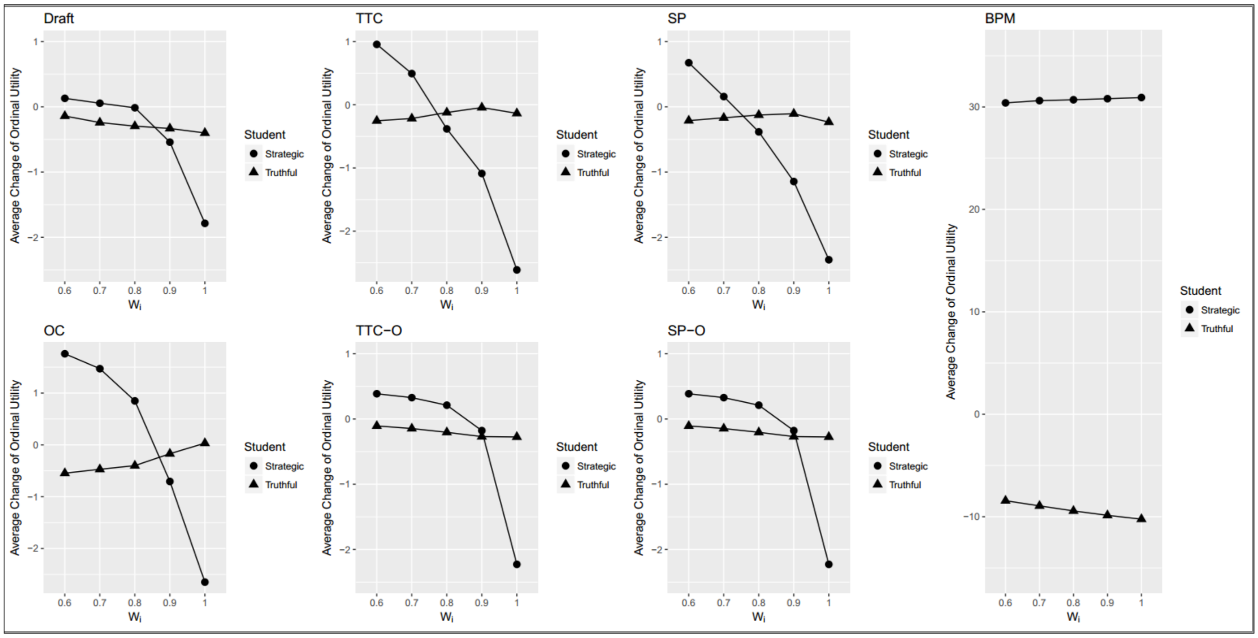


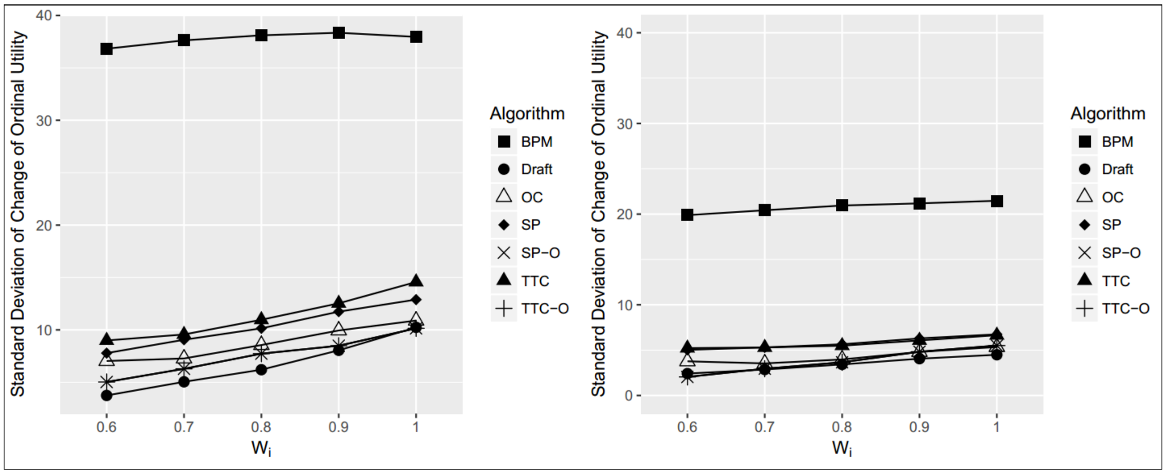
Figure 10 shows that the strategic behavior in BPM hurts truthful students more than in other mechanisms, and that in the manipulated market, the ordinal utility of the truthful students is between 8 units to 10 units below their results in the $P = 0$ market. Moreover, in all mechanisms except BPM, the decrease in ordinal utility among truthful students is less than 0.42. For an alternate set of comparisons, we combine the results of Figures 9 and 10 into Figure 11, placing truthful and strategic students side-by-side for each mechanism.

Figure 11 – Average changes in ordinal utility of strategic and truthful students under $P = 20\%$



This figure shows the gap between changes in ordinal utility of strategic students and truthful students in different mechanisms. As might be expected, under BPM, strategic students significantly gain by misreporting their preferences while their strategic behavior hurts truthful students a bit more than proportionally. By contrast, in TTC-O, SP-O, and the draft mechanisms, the difference between the ordinal utility change of strategic and truthful students is smaller than under other mechanisms. Moreover, among the OC mechanism and other round by round mechanisms, OC shows a larger gap between ordinal utility of truthful students and strategic students. However, this gap is small, below 2.5 units in total rank for OC and less than 1 for all other round-by-round algorithms.

Figure 12 – Left: Standard deviation of change in ordinal utility of strategic students
Right: Standard deviation of change in ordinal utility of truthful students

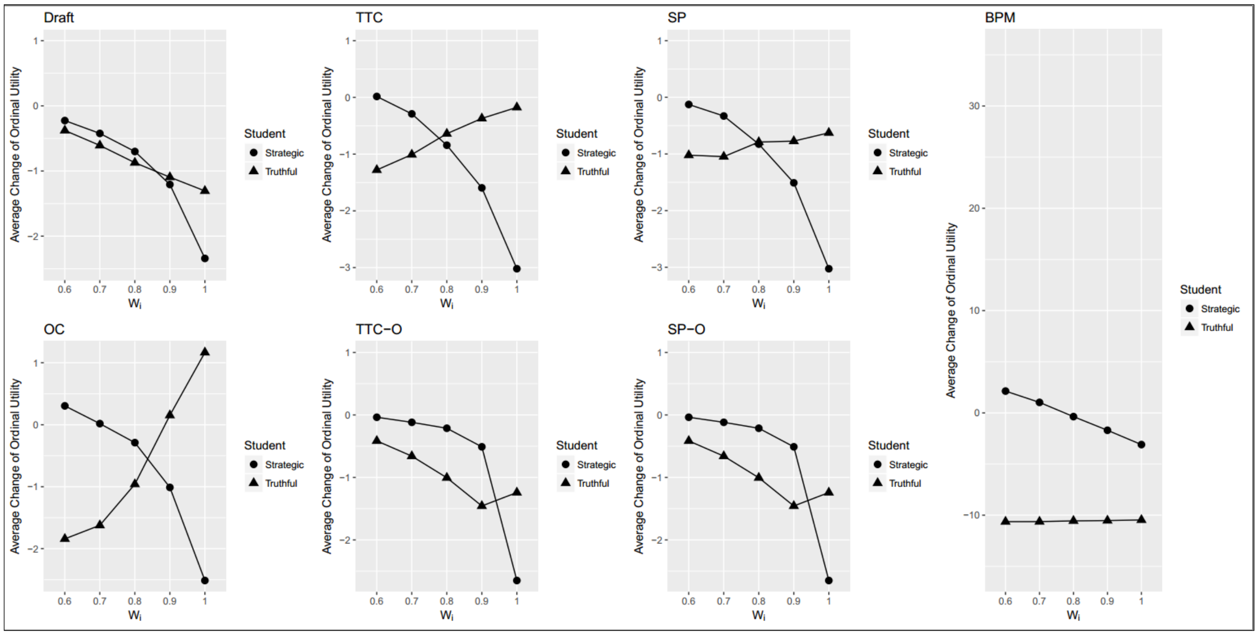


In addition to average effects, the standard deviation of the changes in ordinal utility can be used as a measure of volatility and therefore risk. Though the average effect of strategic deviations may be positive in some scenarios for those who play strategically, a large standard deviation around that mean indicates that some will benefit while others lose utility from their decision to play strategically (though those who benefit will outweigh the losers based on the positive average.) Figure 12 represents the pattern of changes in standard deviation of ordinal utility of strategic and truthful students. In all cases, we see relatively large standard deviations, indicating that deviation from truth-telling is a risky proposition regardless of the mechanism.

5.2. Effect of changing w_i when $P > 20\%$

Here we present an analog to Figure 11 but under $P = 80\%$. Similar results for $P = 40\%, 60\%$, and 100% are given in Appendix B, but the overall story is fairly clear by comparing Figures 11 and 13. As more and more students adopt the non-truth-telling strategy, the benefits of doing so disappear, so that the net effect of a large number of strategic students is purely detrimental.

Figure 13 – Average changes in ordinal utility of strategic and truthful students under $P = 80\%$



In our further results in Appendix B, we see, for example, that when $P = 100\%$ on average all students would be hurt and the average change in ordinal utility is negative for all algorithms and all values of w_i, continuing the trend shown here. When everyone deviates from truth-telling, everyone loses.

5.3. Overall Comparisons of Incentive Compatibility

In §5.2, we considered the average change in ordinal utility (relative to the $P = 0$ case) for strategic students and truthful students. Let these values (averaged over all students in each group for each $(\mathrm{P}, \mathrm{w_i})$ scenario) be denoted $\Delta\mathrm{OU_{str}}$ and $\Delta\mathrm{OU_{tr}}$, respectively. To aggregate the results presented thus far, consider the following metrics. In these three metrics, we did not consider the market with 100% strategic students since there were no truthful students for comparing with strategic students.

Strategic Upside $= \mathrm{average}_{(\mathrm{P},\mathrm{w_i})|\Delta\mathrm{OU_{str}} > \Delta\mathrm{OU_{tr}}}(\Delta OU_{str} - \Delta OU_{tr})$

Strategic Downside $= average_{(P,w_i)|\Delta OU_{str} < \Delta OU_{tr}}(\Delta OU_{tr} - \Delta OU_{str})$

Net Benefit of Strategic Play $= average_{(P,w_i)}(\Delta OU_{str} - \Delta OU_{tr})$

Where the net benefit of strategic play gives a neutral aggregate measure of the average marginal ordinal utility of changing from a truthful student to a strategic student when the scenario is unknown, the upside and downside of strategic play are also informative, in case the downside of a decision is weighted more heavily than the upside (as in loss aversion, see Kahneman and Tversky, 1984).

Figure 14 – Strategic upside (ordinal utility)

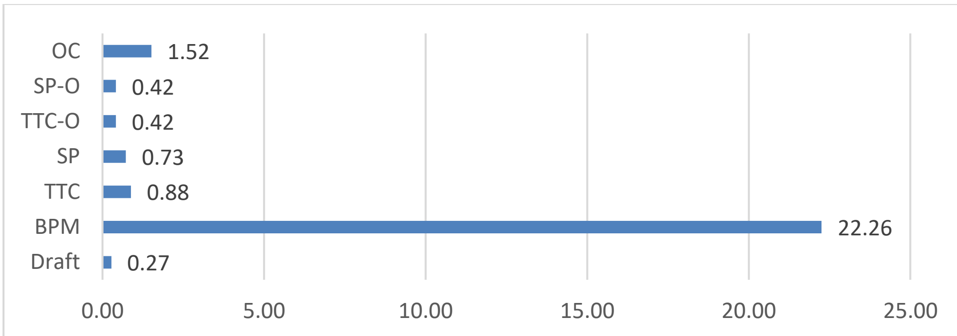


Figure 15 – Strategic downside (ordinal utility)

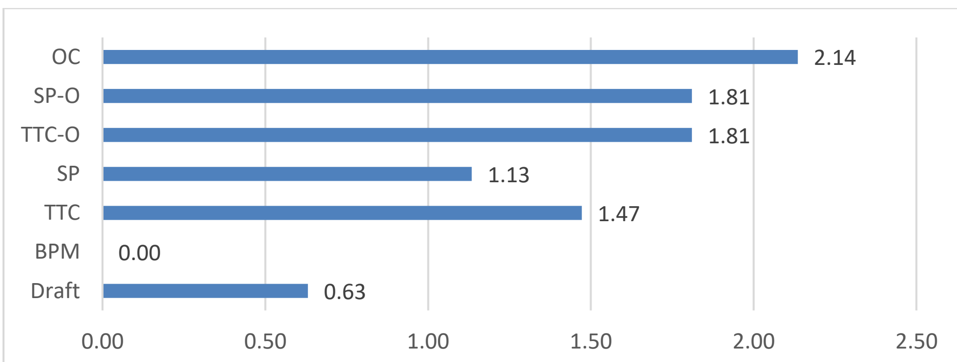
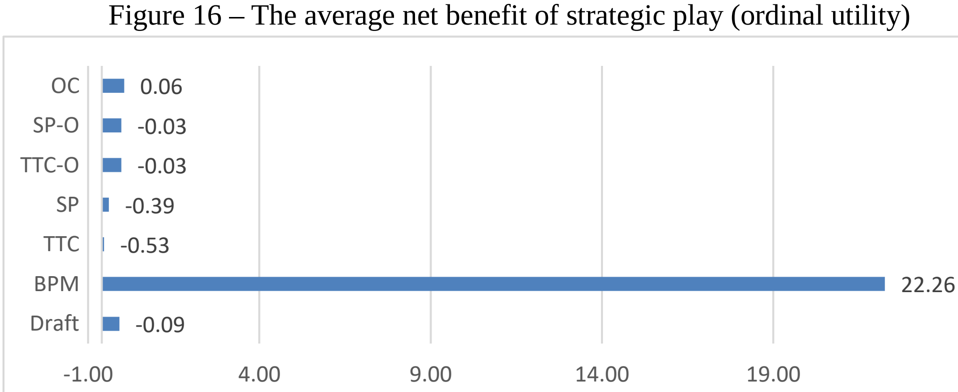


Figure 16 – The average net benefit of strategic play (ordinal utility)



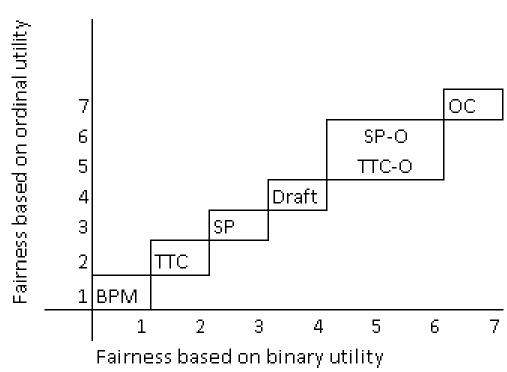
Comparing Figures 14, 15, and 16, we see a few overall conclusions. BPM tends to offer a robust reward to those who play strategically on average, with no downside to doing so. The net benefit to strategic play is negative for all other mechanisms, except for OC, where still it is quite close to zero. Overall, stategic play does not tend to help very much in these latter (non-BPM) mechanisms. For OC, although there is a tiny benefit on average in the net, strategic play is more volatile, introducing the chance of a larger dowside than any other mechanism. The existence of this larger downside risk may in itself be viewed as a deterrent to strategic play.

6. Discussion and Conclusion

This paper has explored five new algorithmic variations to solve the course allocation problem, addressing the relative balance of efficiency, fairness, and incentive compatibility. To understand the performance of these five algorithms, we compared them to the draft and bidding point mechanisms, benchmarks that have been used in practice. We introduced a comprehensive systemization of natural metrics for these objectives, with a novel perspective of analyzing any set of outcomes based on the total (or average), range, and standard deviation of binary, ordinal, and cardinal utility.

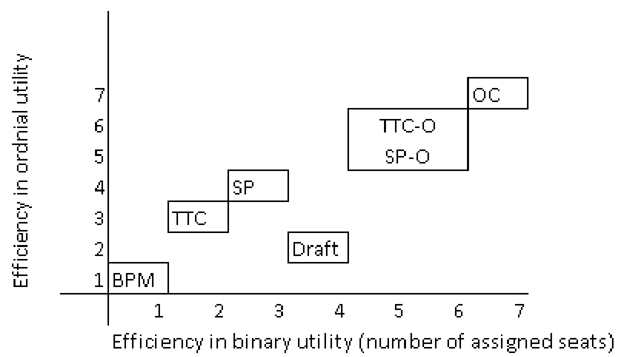
As a final comparison of the mechanisms in question, consider the following selected graphs based on the relative rankings (higher being better) of the seven mechanisms for CAP.

Figure 17 – Ranking the algorithms based on fairness (standard deviation or range)



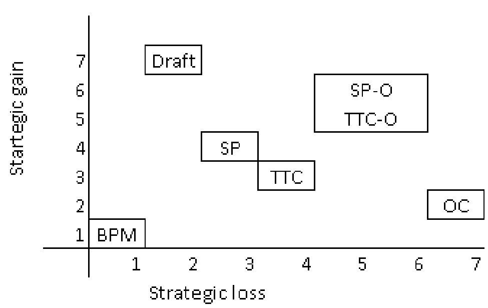
The clearest ordering is in fairness, shown in Figure 17, in which OC dominates, whether one looks at binary or ordinal efficiency. Recall that SP-O and TTC-O switch places with OC when considering cardinal fairness, but with only a small lead, compared to OC's large binary and ordinal lead shown in Figures 2, 4, and 5. Turning our attention to efficiency in ordinal and binary utility, the results remain essentially the same, but with a murkier comparison among SP, TTC, and draft, as in Figure 18.

Figure 18 – Ranking the algorithms based on binary and ordinal efficiency



However, looking at incentive compatibility the results are less clear. In Figure 19, we see that comparisons based on upside opportunities (smaller gains representing less temptation) and downside risks (larger downsides representing a larger deterrent) are less consistent. For example, the draft performs well in not providing large opportunities for upside, while OC performs well in deterring strategic play based on the downside risk of unsuccessful manipulation.

Figure 19 – Ranking the algorithms based on incentive compatibility



Still, perhaps the main conclusion to derive from our experiments on incentives for strategic play is that in the net, except for BPM, all mechanisms provide very little benefit to manipulation (see Figure 16), with the risks of manipulation almost always (at least nearly) outweighing the potential benefits. Indeed, given the nearly universal bad results for BPM, we are inclined to remove it from consideration entirely and consider a comparison among the other mechanisms. The remaining mechanisms perform quite similarly on incentives, leading us to consider a ranking on efficiency and fairness metrics, in which case OC is a clear "winner" among the mechanism investigated here.

If there were reason to believe that incentives were an issue, SP-O or TTC-O seems to provide slightly better incentives than OC, with the smallest degradation on the other metrics overall. In the numerical simulations conducted here, SP-O and TTC-O performed nearly identically, perhaps due to the amount of competition in the large-scale market, though small examples can be devised in which SP-O's use of second prices removes some opportunities for manipulation. This tends to give SP-O an ever-so-slight edge over TTC-O, but one that is nearly negligible. Overall, we hope to have shown that the opportunity for a significant improvement in the performance of course-selection algorithms through the direct application of optimization, whether on a round-by-round or market-wide level.

Finally, our additional experiments shown in Appendix D indicate that SP-O and TTC-O maintain better fairness properties in cases of extreme correlation, as opposed to OC which will instead sacrifice fairness for efficiency in such circumstances. However, these extreme cases seem less likely in practice, and our results indicate very little hope for any mechanism in such cases. If everyone is in complete agreement, trying to use a mechanism to get them to admit where their preferences are different is futile. Still, these additional results point to the possibility of strengthening the OC algorithm through additional fairness constraints and reinforce the necessity of considering round-by-round mechanisms like SP-O and TTC-O in discussions of future applications.

References

Akbarpour, M., Nikzad, A., (2015). Approximate random allocation mechanisms, working paper

Budish, E., Cantillon, E., (2012), The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard, the American Economic Review. Vol 102, no 5.

Budish, E., (2011) The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes, Journal of Political Economy Vol. 119, 1061-1103.

Budish, E., Che, Y., Kojima, F., Milgrom, P., (2013). Designing Random Allocation Mechanisms: Theory and Applications. American Economic Review, 103(2): 585-623

Diebold, F., Aziz, H., Bichler, M., Matthes, F., Schneider, A. (2014). Course allocation via stable matching. Business. Information System Engineering 6 (2), 97–110

Ehler, L., & Klaus, B, (2003), Coalitional strategy-proof and resource-monotonic solutions for multiple assignment problems, Social choice and welfare, Vol. 21, 265-280.

Hatfield, J., (2009), Strategy-proof, efficient, and nonbossy quota allocations, Social Choice Welfare. Vol 33.

Kahneman, D. and Tversky, A., (1984). Choices, Values, and Frames. American Psychologist 39 (4): 341–350.

Kominers, S.D., Ruberry, M., Ullman, J., (2010). Course Allocation by Proxy Auction, Internet and Network Economics. 551-558.

Krishna, A., & Ünver, U., (2008), Improving the efficiency of course bidding at business schools: Field and laboratory studies. Marketing Science, Vol.27, 262-282.

Morrill, T., (2012), Making Efficient School Assignment Fairer.

Nguyen, T., Peivandi, A., Vohra, R., (2014). One-sided matching with limited complementarities. working paper.

Nogareda , A., Camacho, D. (2015). Optimizing Satisfaction in a Multi-Courses Allocation Problem. Intelligent Distributed Computing IX. 616, 247-256

Papai, S., (2001), Strategyproof and nonbossy multiple assignments, Journal of public economic theory, Vol. 3, 257-271.

Roth, A., Sotomayor, M., (1990). Two-Sided Matching. Cambridge: Cambridge University Press.

Schummer, J., Vohra, R. V. (2007). Mechanism design without money. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani (Eds.), Algorithmic Game Theory (pp. 243-265). Cambridge: Cambridge University Press.

Sönmez, T. & Ünver, U., (2010). Course bidding at business schools, International Economic Review, Vol. 51, 99-123.
http://www.hbs.edu/mba/registrar/crossregistration/Documents/fsched13.pdf
http://www.hbs.edu/mba/registrar/crossregistration/Documents/wsched14.pdf

Appendix A. Market simulation details for the efficiency and the fairness results

We assume that the value of each course for each student is the combination of a field-specific value for course belonging to the field (or major plan of study), plus an independent course value for the student. This setup tends toward correlated preferences among students according to their interest in a field/major denoted f. Therefore, a value $V_{ij'} = \sum_{[\forall f:j' \in f]} \alpha_{fj'} + \beta_{ij'}$ is calculated, where the value $V_{ij'}$ of course j' (considering all of its sections) for student i is defined equal to $\alpha_{fj'}$ which is the random order of field f for student i multiplied by 5 plus $\beta_{ij'}$, a random number drawn from the normal distribution ($\mu = 10, \sigma = 10$).

Then, the top M_1 courses with highest $V_{ij'}$ are selected from 83 courses for each student. M_1 is independently chosen for each student from the discrete uniform distribution on the interval [30, 35]. To generate correlated yet distinct values for particular sections of the same course, we generated random normal number $CS_{ilj'}$ for the l^{th} section of each selected course j' of student i, and set the section value to $VS_{ilj'} = V_{ij'} + CS_{ilj'}$. For each student, all course-sections are sorted based on $VS_{ilj'}$ and the top $M1$ course-sections are selected. Re-indexing course sections to a single list translates the l^{th} section of each selected course j' to the index, and thus $VS_{ij} = VS_{ilj'}$. The ordinal values of course r_{ij} are the inferred based on VS_{ij} to span the integers between $[100 - M_1 + 1, 100]$. Therefore, r_{ij} for student i's top course is 100, for her second favorite course is 99, and so on.

Finally, the u_{ij} values of course-sections of each student are assigned as follows. Depending on the number of courses evaluated by a student, approximately 10% (and then 40%, 40%, 10%, respectively) of the courses receive a draw from a discrete uniform distribution [0,10] (and then [10,50], [50,100] and [100,200], respectively). For each student, these values are normalized to sum up to 1000 and then monotonically assigned as u_{ij} values in the same order as the already assigned r_{ij} values determined above.

Appendix B. Further incentive compatibility results when $\boldsymbol{w_i}$ for truthful students is equal to 0, 0.25, 0.5, and 0.75, and when $P > 20\%$.

In the first part of this appendix, we assume that w_i of truthful students is 0, 0.25, 0.5, and 0.75 and show measurements similar to those shown in §5.3.

Figure B.1 – Strategic upside when w_i for truthful students is 0

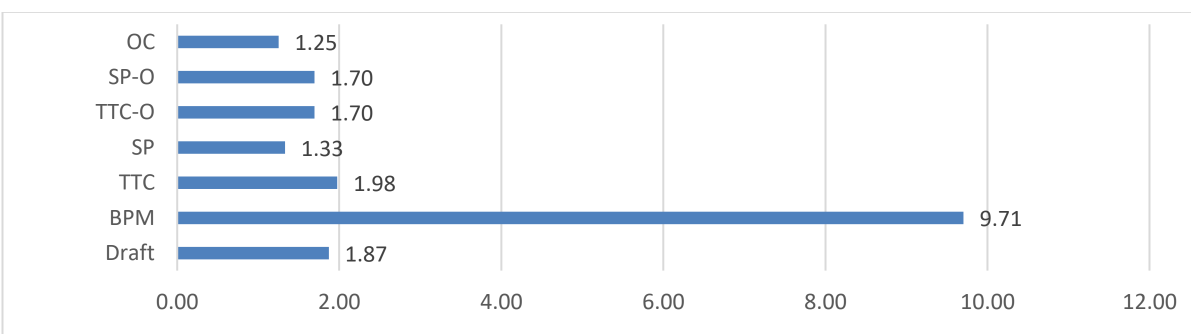


Figure B.2 – Strategic upside when w_i for truthful students is 0.25

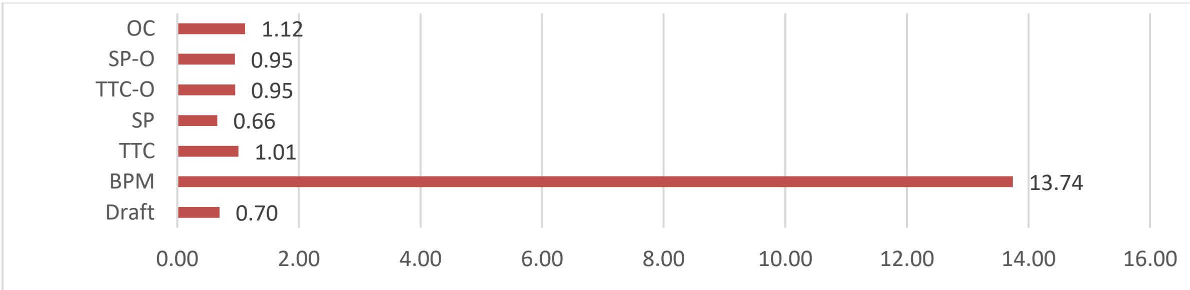


Figure B.3 – Strategic upside when w_i for truthful students is 0.5

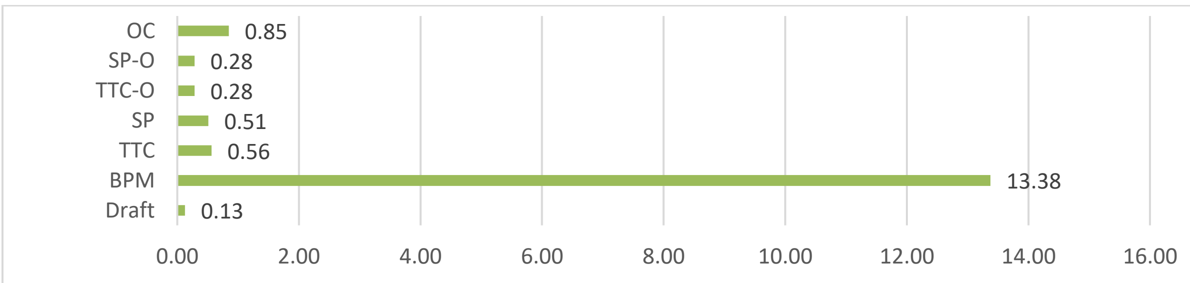


Figure B.4 – Strategic upside when w_i for truthful students is 0.75

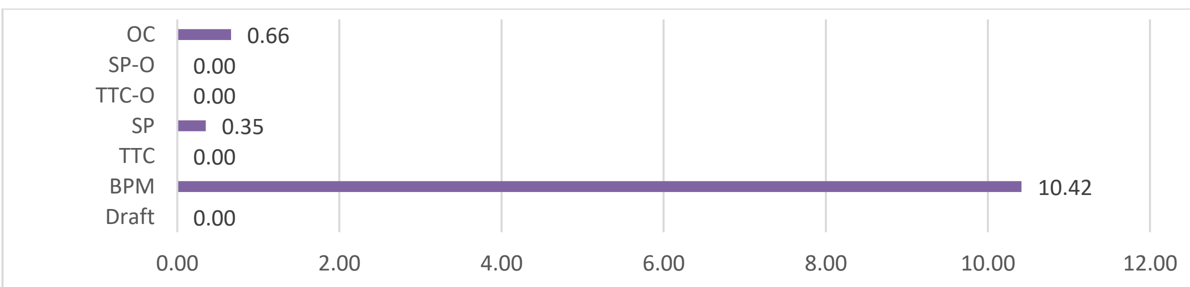


Figure B.5 – Strategic downside when w_i for truthful students is zero

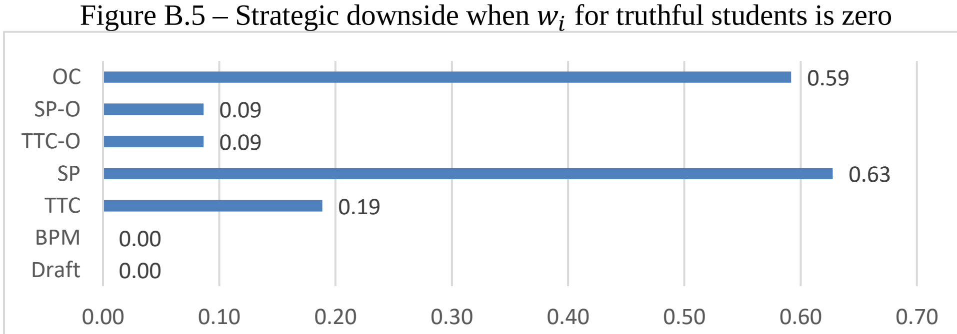


Figure B.6 – Strategic Downside when w_i for truthful students is 0.25

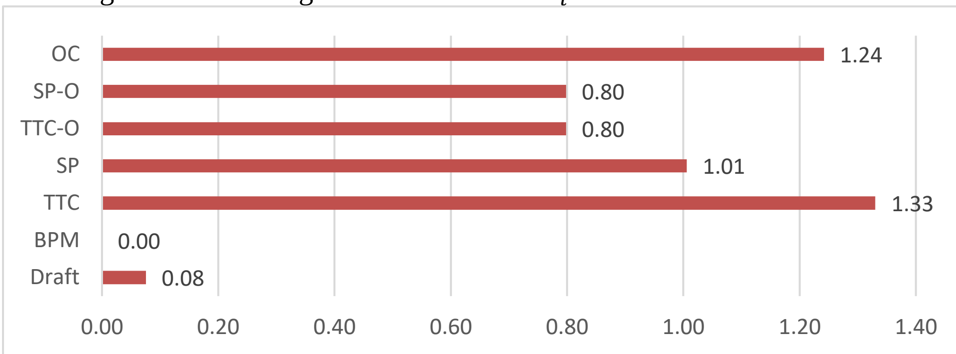


Figure B.7 – The Strategic Downside when w_i for truthful students is 0.5

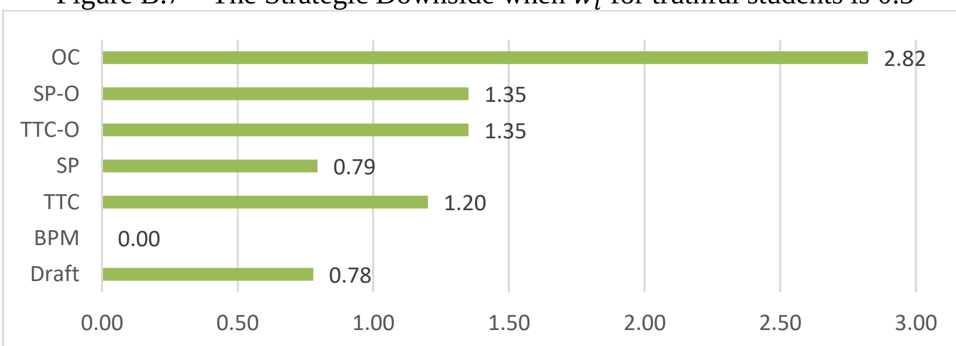


Figure B.8 – The Strategic Downside when w_i for truthful students is 0.75

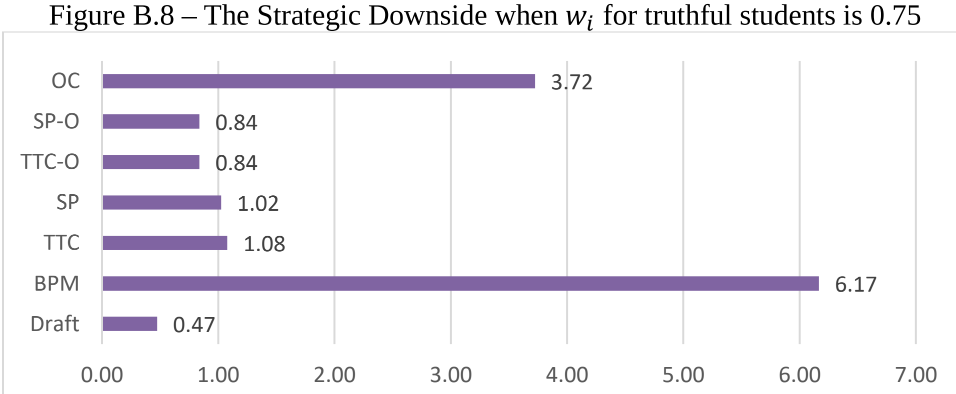


Figure B.9 – The Net Benefit when w_i for truthful students is zero

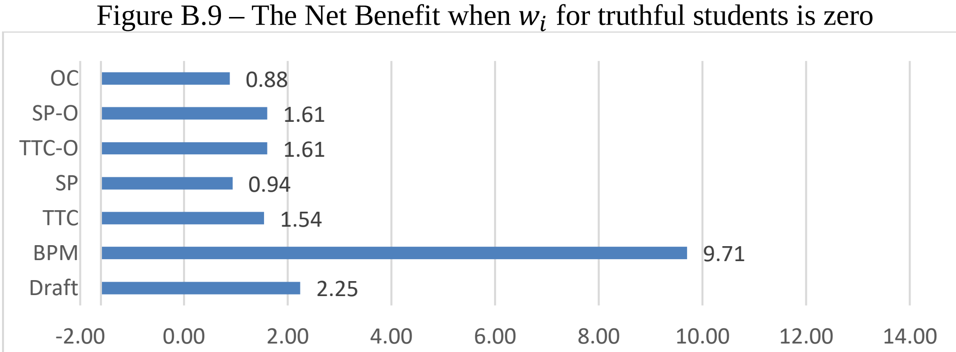


Figure B.10 – The Net Benefit when w_i for truthful students is 0.25

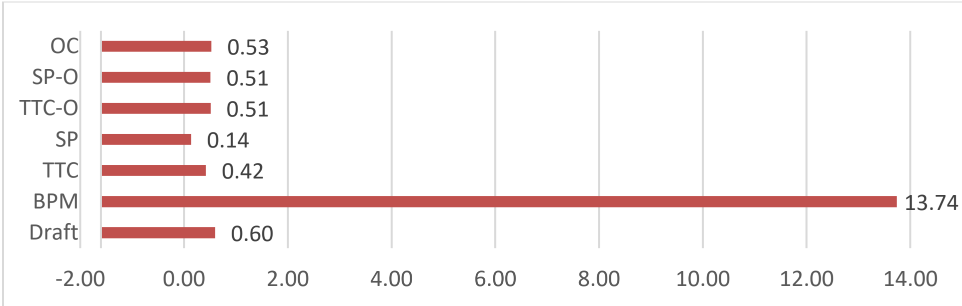


Figure B.11 – The Net Benefit when w_i for truthful students is 0.5

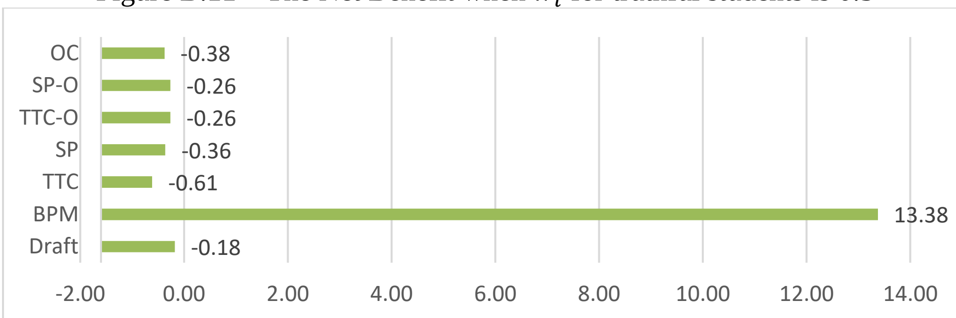
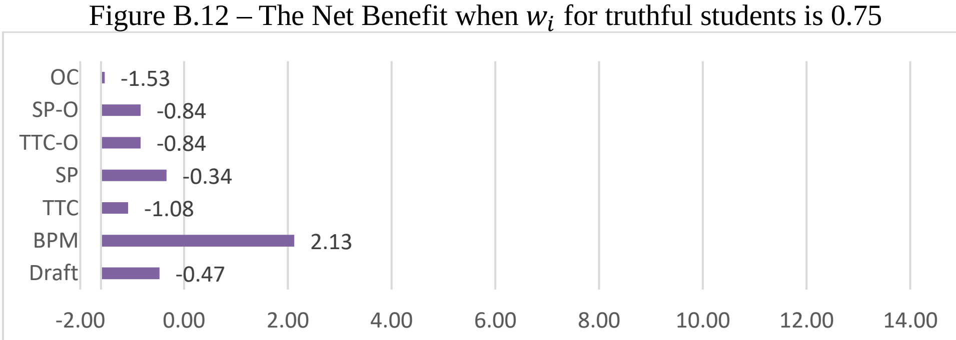


Figure B.12 – The Net Benefit when w_i for truthful students is 0.75



For the second part of this appendix, we provide additional results for the main setup (i.e., when w_i for truthful students is 0.5) but when the proportion of strategic students, P, varies.

Figure B.13 – Average changes in ordinal utility of strategic and truthful students under $P = 40\%$

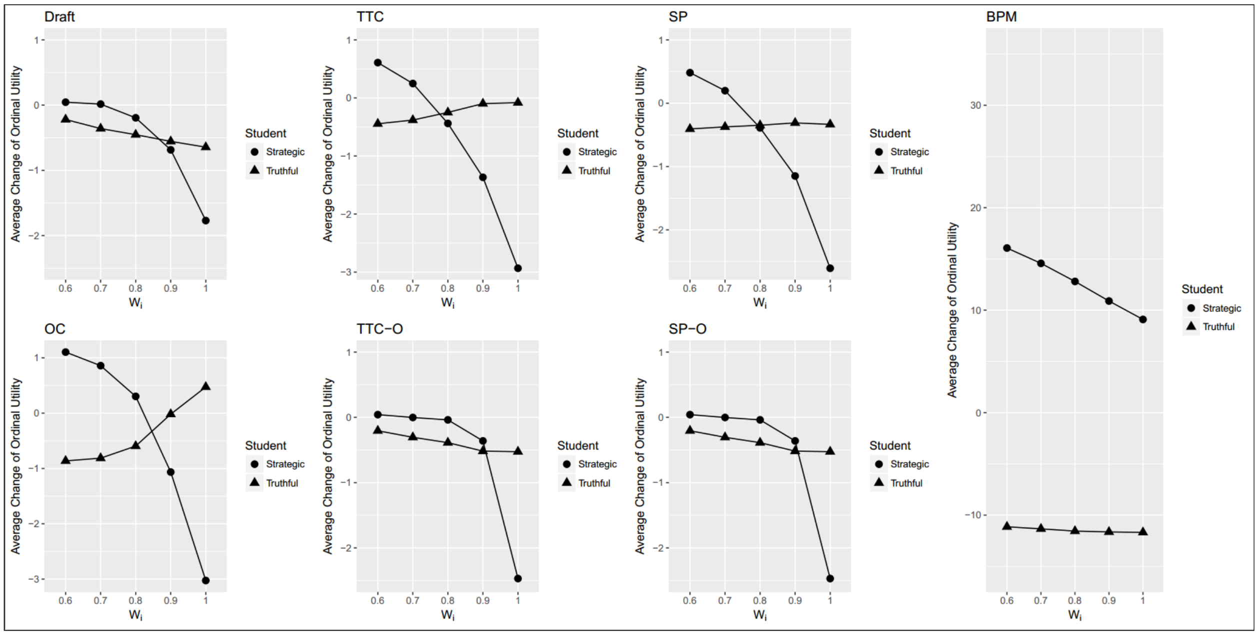


Figure B.14 – Average changes in ordinal utility of strategic and truthful students under $P = 60\%$

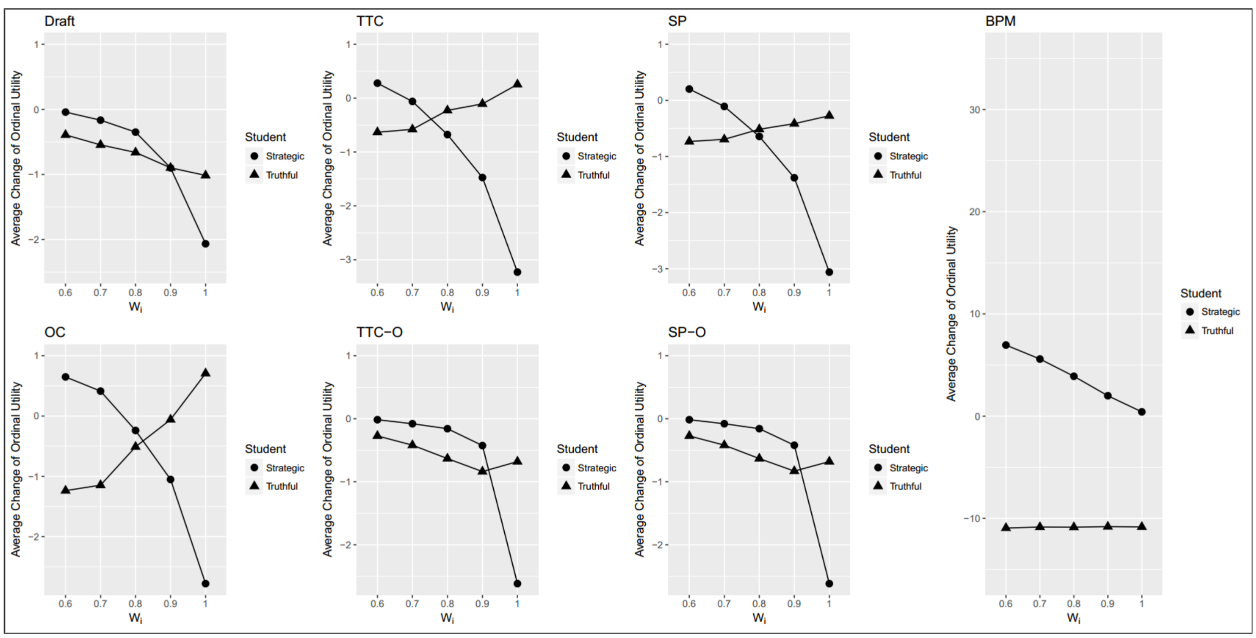


Figure B.15 – Average changes in ordinal utility of strategic and truthful students under $P = 100\%$

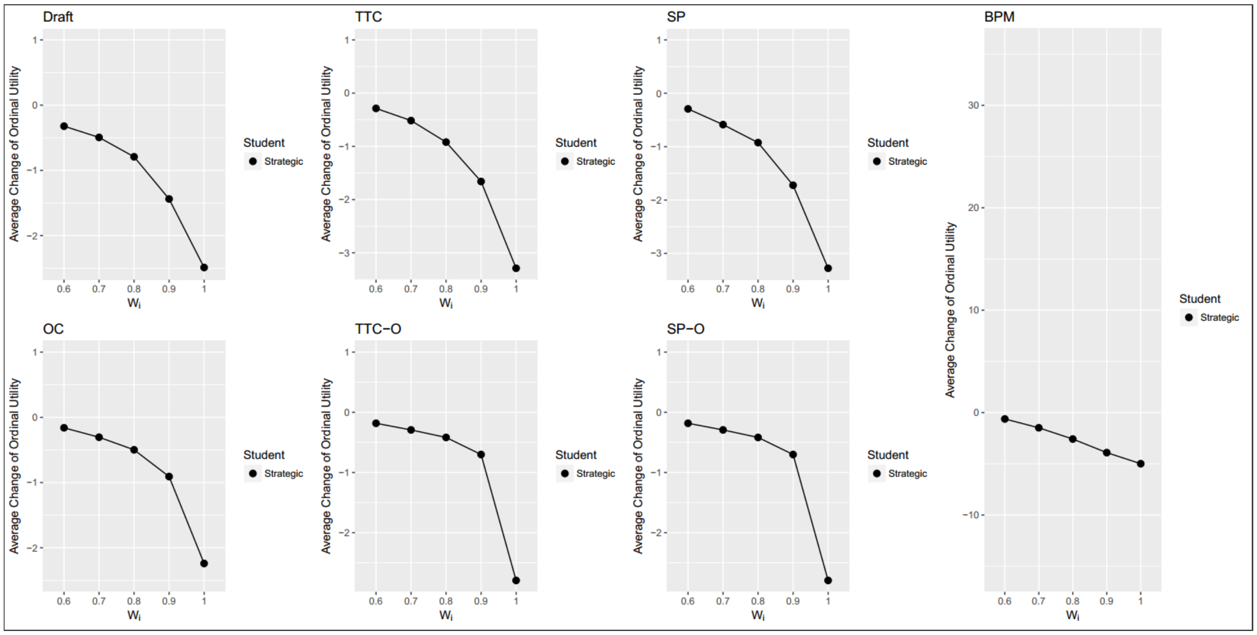
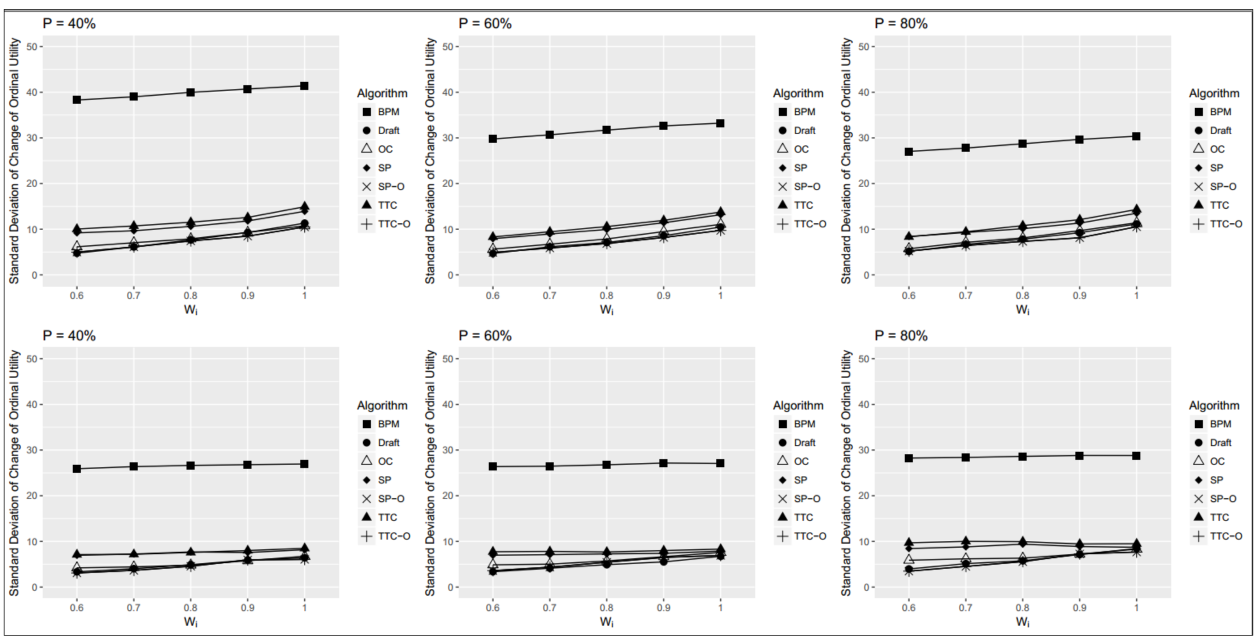


Figure B.16 – Top – Standard deviation of change in ordinal utility of strategic students with different P. Bottom – Standard deviation of change in ordinal utility of truthful students with different P.



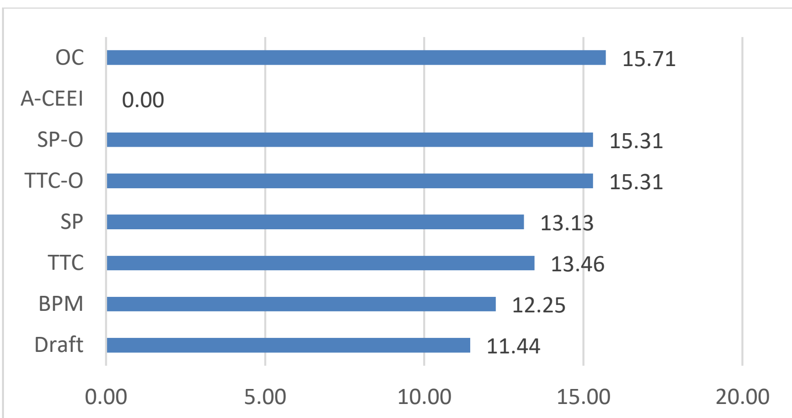
Appendix C. Comparison to the A-CEEI mechanism.

To compare the A-CEEI algorithm with the algorithms described here, we first found that the computational time burden of running that algorithm on the same size problems discussed in this paper was prohibitive. So we instead used the smaller setup of Othman et. al. (2011), but still, based on computation time we had to limit ourselves to just five market instances. Each of these markets includes $n = 250$ students and $m = 50$ courses. Each course has only one section and the possibility of course overlap was not considered. Also, each course has $q_j = 27$ seats and each student needs $k = 5$ courses. In each of these five cases, it was possible to run *all seven* algorithms described in the earlier figures of this paper in under 6 seconds, while a Tabu search implementation of A-CEEI consistently reached our ***2-hour*** time limit without converging to a market clearing error below the defined threshold. Budish (2011) proved that the market clearing error threshold for each market is $\sqrt{(k \times M)/2}$ which is equal to 15 in our experiments, though we could not achieve values lower than about 49.5 on average. This metric of infeasibility translated to an average of 119.4 seats assigned beyond the capacity (in total over several highly demanded courses) while simultaneously resulting in several courses with empty seats, totaling 131.2 unused seats on average (because their price was set too high.)

In short, we find quite poor performance for the A-CEEI method compared to the others discussed here. All algorithms assigned the max possible number of seats (50*5=1250 seats) to students except A-CEEI which assigned on average 12 seats fewer than other algorithms. Also, in all random trials, all algorithms except A-CEEI assigned exactly 5 courses to each student while in A-CEEI on average the worst student gets 2.2 courses. Accordingly, the average range of binary utility for all algorithms is zero while it is 2.8 for A-CEEI algorithm. Similarly, the average standard deviation of binary utility for all algorithms is zero while for A-CEEI it is 0.315.

Further results are provided as follows in Figures C.1 through C.6:

Figure C.1 - The average ordinal utility per student per algorithm, showing amounts above A-CEEI benchmark with average ordinal utility of 432.34 per student.



Because the Othman et al. (2011) setup is more competitive market, with total supply (50 * 27 = 1350 seats) not much more than total demand (250 * 5 = 1250), in most algorithm runs no student is awarded her best bundle. Here, the ordinal value of the best bundle is 490 points ($= 100 + 99 + 98 + 97 + 96$).

Figure C.2 - The average range of ordinal utility per student per algorithm

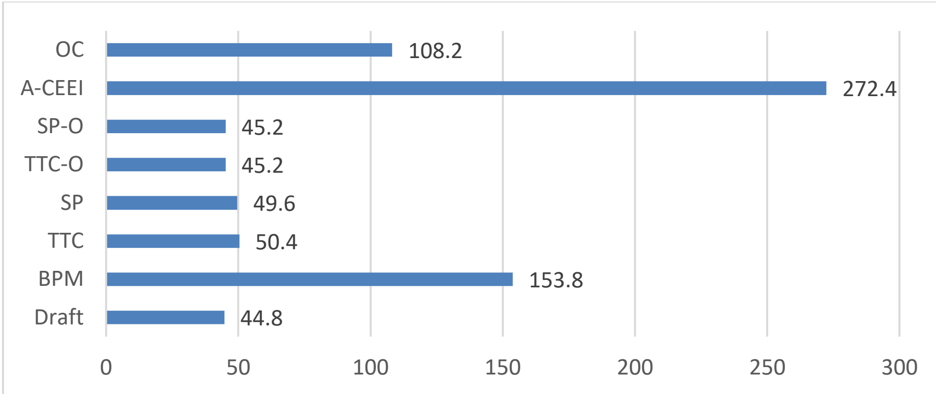


Figure C.3 – Average standard deviation of ordinal utility per student per algorithm

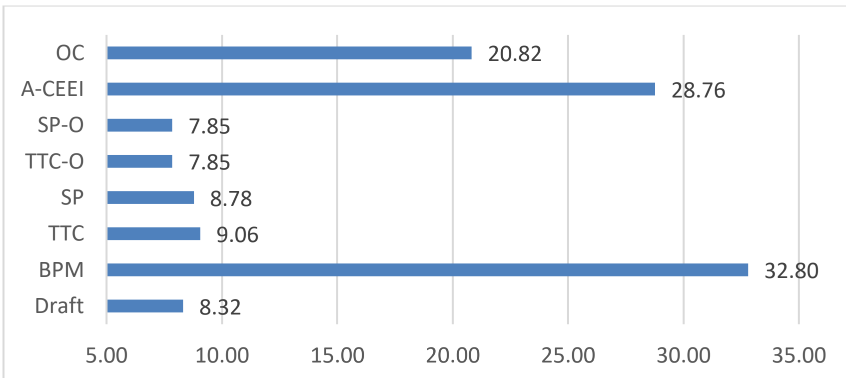


Figure C.4 – The average cardinal utility per algorithm, amounts above A-CEEI benchmark with average cardinal utility 156.73

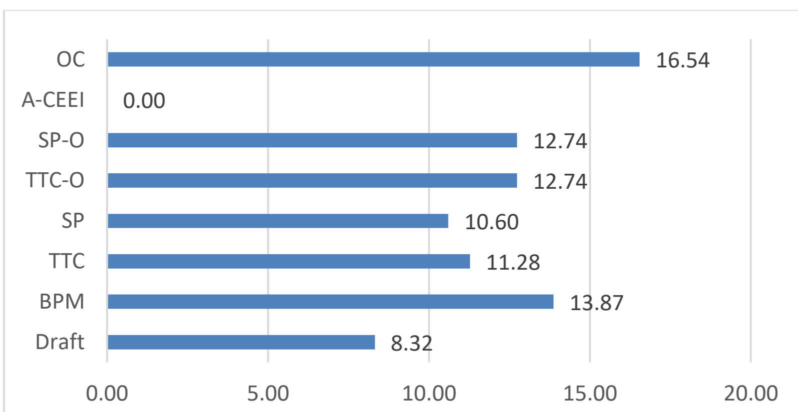


Figure C.5 – The average range of cardinal utility per algorithm

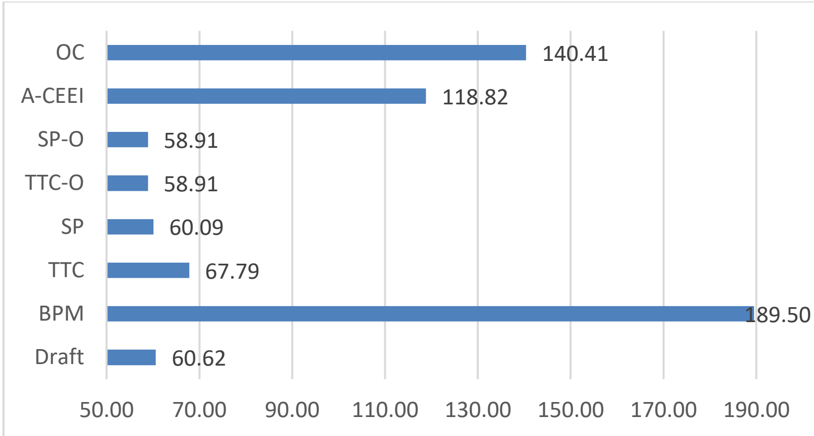
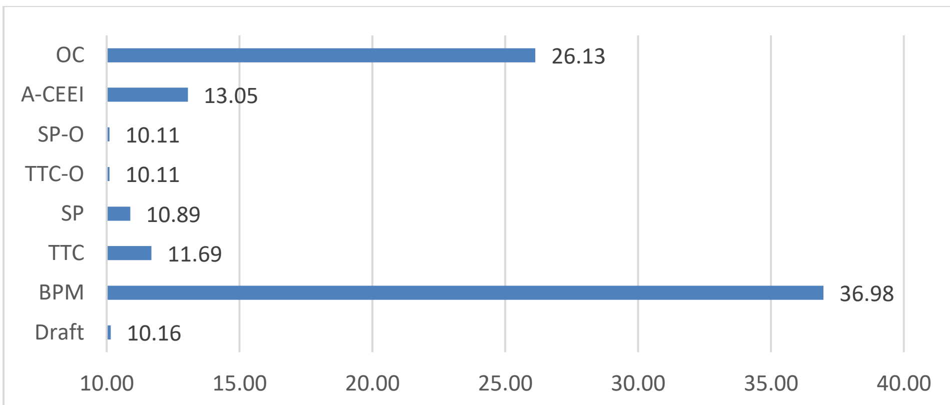


Figure C.6 – Average standard deviation of cardinal utility per algorithm



Appendix D. Comparisons on Random Instances, Varying the Common-Value Weighting for True Values.

In a final set of experiments, we followed the simpler simulation structure of Kominers et. al. (2010) to test the performance of the investigated algorithms in extreme cases. Similar to their model, each course only has one section and no field of study was defined. However, unlike their model, we maintained our assumptions of overlapping courses and that a student will not ranks all possible courses. Accordingly, to quickly view the performance of the mechanisms under varying degrees of common-value, equation (21) was used to define the true value of course j for student i.

$$u_{ij} = (1 - W) * v_{ij} + W * g_j \tag{21}$$

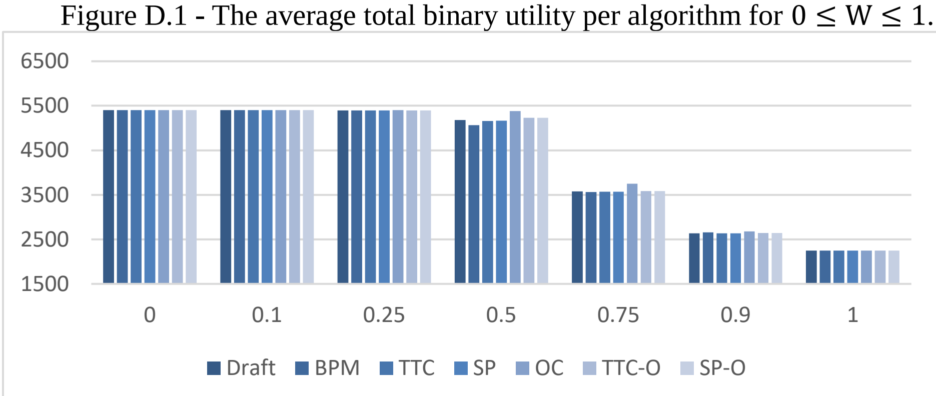
In equation (21), v_{ij} and g_j are independently selected from the standard normal distribution and W was varied as a treatment parameter. When $W = 1$, all students exactly bid equally for each course, while if $W = 0$ students utilities for each course are totally independent. When $0 < W < 1$, student preferences are partially correlated in varying degrees. In our analysis, we compared algorithms in five conditions: W equals to 1, 0.9, 0.5, 0.1, and 0.0. As W increases, the correlation among student preferences increases and bid values and rank values of each course among students become more similar to each other.

Incentives were not considered in this set of experiments. Where §5 and Appendix C varied a personalized parameter w_i (in equation 20) to capture manipulation of true values for a studied of incentives, here the single value W is varied to change true preferences across the entire market. So here, rather than asking how manipulation affects outcome when some pretend to have a larger than truthful amount of common value, we ask how does the market as a whole perform when more or less common value is present in true valuations. For a further element of robustness, we drop some of the constructs of realism described in Appendix A, in which students have major fields of study. Here the focus is instead on more random instances, trying to see how well our algorithms perform under more extreme circumstances.

D.1 Binary Utility

Looking at Figure D.1, our results show that when W is small (0 or 0.1) and student preferences over their top 30 to 35 courses are mostly random and uncorrelated to each other, each student did receive exactly 6 courses in all algorithms and in all 100 market instances therefore all algorithms could assign 5400 seats to students. When W is equal to 0.5, the only algorithm which assigns all 5400 seats is OC while BPM misses the most seats in comparison to other algorithms, as before. When W increases to 0.9, the Draft mechanism assigns the lowest number of seats among different algorithms. Without ranking all courses (students only list their favorite 30-35 courses out of 83 available) assigning all seats becomes increasingly difficult for all algorithms as W increases. As everyone's preferences converge to the same favorite courses, it becomes impossible to assign courses that no one has asked for. As seen in Figure D.1, OC still achieves the best binary utility outcome, assigning the maximum number of seats among different algorithms. However, even OC cannot assign more than about half of offered seats as W approaches 1.

Figure D.1 - The average total binary utility per algorithm for $0 \leq W \leq 1$.

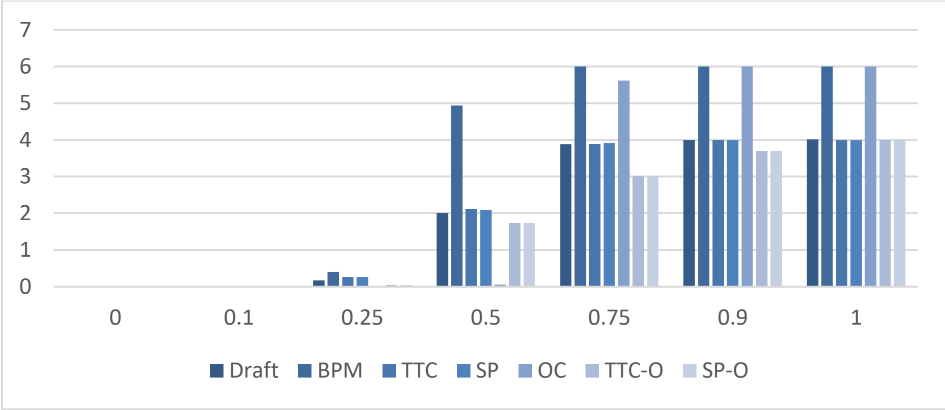


D.2 Range of Binary Utility

As already noted, for small values of W such as 0 or 0.1 is exactly 0 since in these conditions all algorithms are able to assign 5400 seats to students and all students can take 6 courses. However, as W increases not all algorithms are able to assign the maximum number of courses. When W is equal to 0.5, BPM achieves the highest range of binary utility and the most unfair solutions. These results reveal that in 99 instances of 100 instances BPM assigned 2 or fewer seats to the student with minimum number of courses while for OC this number is always 6. However, when W increases to 0.9 or 1, in each of our simulated market instances, there is at least one student who could not receive even one course based on OC algorithm. In these examples, the rank and bid of each specific course for all students is almost the same and the algorithm does not care to assign a specific course to a student who already has other courses or to a student with no course. As a result, when preferences are so correlated that less than full schedules are necessary, OC has no built-in safe guard to assign courses equitably. Thus, if OC were to be implemented in situations of extreme correlation additional constraints would be needed. In other words, when W is large, we need to add constraints to OC to keep the number of courses among students similar in order to achieve similar results as in the more realistic (major-field of study based) simulations.

However, we expect not to face this issue in SP-O and TTC-O algorithms, based on the round-by-round methods' explicit goal of equity. Our results confirm this expectation (see Figure D.2), and we see that when all preferences are (nearly) identical SP-O can achieve the fairest results and can be considered as the most reliable algorithm in achieving fair solutions.

Figure D.2 - The average range of binary utility per algorithm for 0 ≤ W ≤ 1.



D.3 Ordinal and Cardinal Utility

When W is small and student preferences are mostly uncorrelated to each other, again we find all algorithms performing similarly in average ordinal and cardinal utility per person with slightly better results for our proposed algorithms. As expected, OC which first maximizes total market ordinal utility always performs best on the average ordinal performance. When W increases to 0.5, OC and then TTC-O and SP-O achieve the most efficient solutions based on ordinal and cardinal utility while BPM is the most inefficient algorithm based on these metrics. When W becomes close to 1, all algorithms perform similarly based on average ordinal and cardinal utility, as seen in Figures D.3 and D.4.

Figure D.3 - The average ordinal utility per student per algorithm for 0 ≤ W ≤ 1.

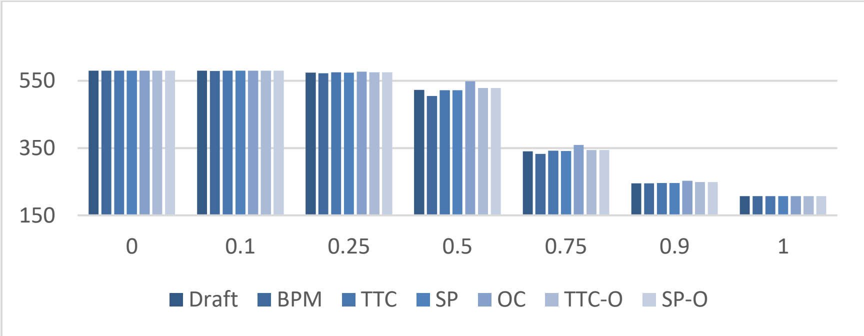
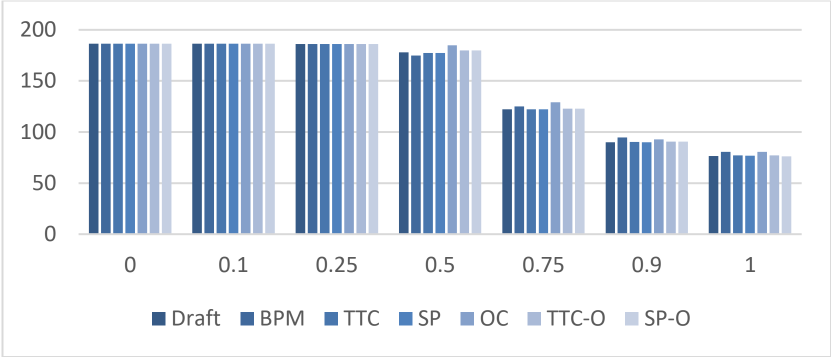


Figure D.4 - The average cardinal utility per student per algorithm for 0 ≤ W ≤ 1.



D.4 Range and Standard Deviation of Ordinal and Cardinal Utility

Our final set of results are shown in Figure D.5 through D.8, concerning the range and standard deviation as fairness metrics on utility. Where OC again dominates at the intermediate point in which $W = 0.5$, consistent with the main results of the paper, these fairness results tend to vanish as the correlation grows. In its focus on achieving good ordinal and cardinal market performance, OC will randomly favor some students; with (nearly) identical students some must win and some must receive very poor (even empty) schedules in the best overall outcome. Still TTC-O and SP-O show that with only mildly worse overall performance, these ranges and standard deviations can be kept lower. These latter mechanisms are therefore seen to be more robust to extreme correlation with respect to fairness, converging toward a pure rationing outcome as students get closer to identical preferences.

Figure D.5 – The average range of ordinal utility per algorithm for $0 \leq W \leq 1$.

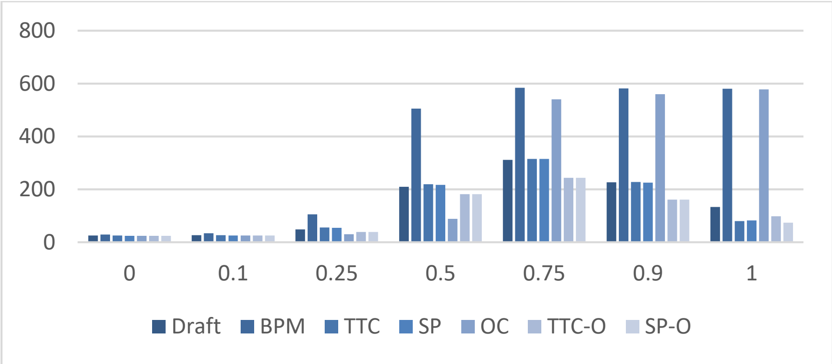


Figure D.6 – The average range of cardinal utility per algorithm for $0 \leq W \leq 1$.

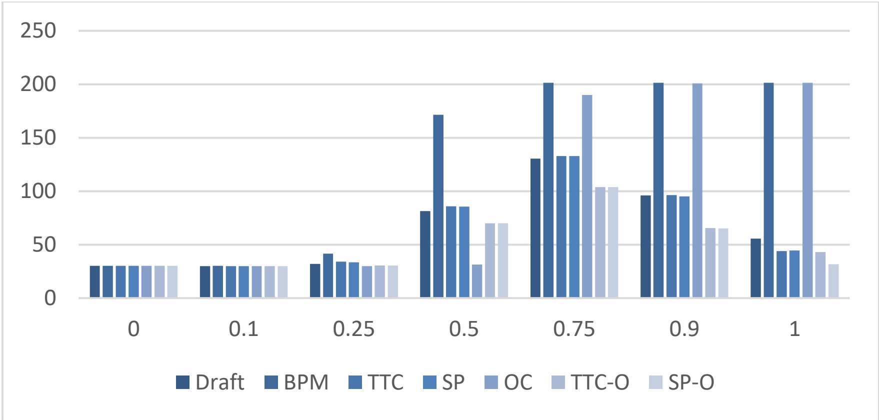


Figure D.7- The standard deviation of ordinal utility per algorithm for $0 \leq W \leq 1$.

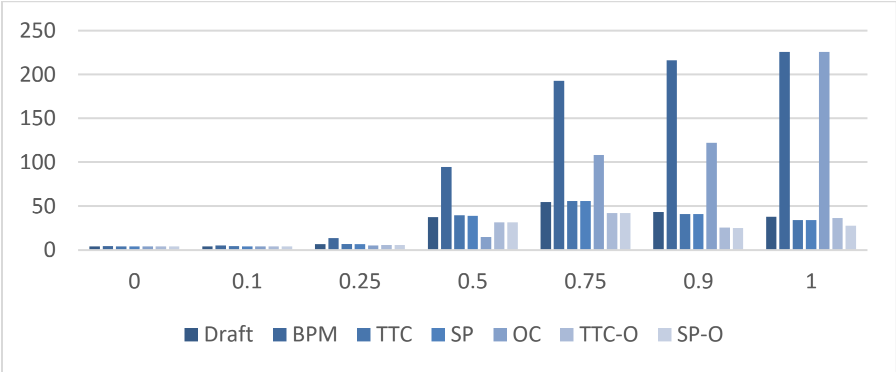
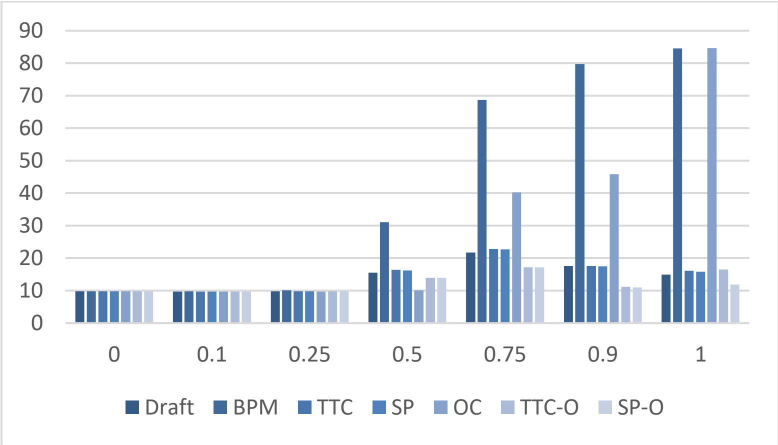


Figure D.8- The standard deviation of cardinal utility per algorithm for $0 \leq W \leq 1$.



Chapter 3. On Finding Stable and Efficient Solutions for the Team Formation Problem

Abstract:

The assignment of personnel to teams is a fundamental and ubiquitous managerial function, typically involving several objectives and a variety of idiosyncratic practical constraints. Despite the prevalence of this task in practice, the process is seldom approached as a precise optimization problem over the reported preferences of all agents. This is due in part to the underlying computational complexity that occurs when quadratic (i.e., intra-team interpersonal) interactions are taken into consideration, and also due to game-theoretic considerations, when those taking part in the process are self-interested agents. Variants of this fundamental decision problem arise in a number of settings, including, for example, human resources and project management, military platooning, sports-league management, ride sharing, data clustering, and in assigning students to group projects. In this paper, we study a mathematical-programming approach to \team formation" focused on the interplay between two of the most common objectives considered in the related literature: economic efficiency (i.e., the maximization of social welfare) and game-theoretic stability (e.g., finding a core solution when one exists). With a weighted objective across these two goals, the problem is modeled as a bi-level binary optimization problem, and transformed into a single-level, exponentially sized binary integer program. We then devise a branch-cut-and-price algorithm and demonstrate its efficacy through an extensive set of simulations, with favorable comparisons to other algorithms from the literature.

1. Introduction

Consider the task of assigning n workers to m teams of equal size (assume that m evenly divides n). Each worker scores each co-worker, being asked to express an integer from 1 to (say) 10, with 10 being a most preferred teammate.

Through the application of constraints, it is easy to imagine many practical variations of this basic setting. For example, perhaps there are *m* project managers among the *n* workers, and each team must have exactly one project manager. Each project manager is tied to a specific project plan, and workers consider the two together in their scoring. Variations with more constraints (e.g., at least one marketing person, at least two from IT, etc.) abound.

This process of putting people into teams and observing their satisfaction (or lack thereof) happens incredibly frequently in practice, yet obtaining maximum-score solutions for even modestly sized instances can be computationally prohibitive. Indeed, the maximization of the sum of the interpersonal scores among teammates is NP-Hard, and is too difficult to solve for even a few dozen workers, depending on how sophisticated of a formulation is used.

Beyond utility maximization, the economic concept of stability has played a major role in similar types of preference-reporting mechanisms. In many games, we ask if a subset of players could all improve their situation by breaking away from the game and getting together by themselves. Such a situation is called *unstable*, and is considered bad for morale or for the survivability of the game as an institution. Further, it encourages the subset to collusively misstate their preferences to achieve the better outcome. (By reporting maximum scores for those in the subset, and zero for all others, group manipulation may often prove beneficial, undermining the goals of the mechanism.) Some job markets (modeled as marriage markets, such as medical-residency matching) have been shown in the classical literature to always admit stable solutions, sometimes many. Other simple markets may have no stable solution, such as in the *stable roommate problem* (which can be modeled as our current problem with team-size $s = \frac{n}{m} = 2$ and no constraints besides $s = 2$).

This leads us to consider the problem of finding a solution that minimizes the amount that any subset (or *coalition*) of players could mutually improve their situation by breaking away and forming a new valid team. We call this the *maximum uplift coalition*, and consider the minimization of the maximum uplift.

When the minimum is zero, the solution is stable, but even when no stable solution exists, the minimum provides a measure of instability.

A restriction to equal-sized teams is used throughout this paper; other constraints based on worker characteristics may or may not be included. Since the case when m does not evenly divide n but team sizes as close as possible to uniform are required can be modeled by adding dummy workers and a constraint enforcing that at most one dummy worker is assigned per team, we assume m evenly divides n throughout for simplicity. This (nearly) equal-size team assumption is crucial, driving the interesting behavior of this type of system (in contrast to hedonic games). With equal-size teams, individual actions have less influence, making group uplift a more interesting consideration than unilateral deviation.

After a broad review of the literature (§2), we provide a descriptive model (§3) and then refine to a more usable single-level reformulation (§4). We provide and then compare three algorithmic implementations (§5 and §6) before comparing to benchmark heuristics from the literature (§7) across four distinct preference models.

2. Related Literature

2.1. Team Dynamics

The study of interpersonal dynamics in teams has a huge stream of literature within the management and organizational psychology community. Though too deep to expound upon broadly here, Gardner et al. (2017), for example, have considered formal models with quadratic objective terms based on pairwise utility realized among those individuals placed in the same group. See Mathieu et al. (2015) for a survey of the dynamics of people working in teams and their influence on team formation. One perspective is that a team planner will often have several practical constraints on acceptable teams, for example that each team must include a minimum number of people from a certain gender, ethnic group, or having a specific skill set (Campion et al. 1993). Moreover, several models and algorithms have been developed to solve team partitioning problems that consider different forms of skill constraints (Farhadi et al. 2011, Gutierrez et al. 2016, Chen et al. 2012, Agrawal et al. 2014). These provide support for our constraint-based architecture, with each team needing to satisfy quotas or caps on workers of particular skill-sets or other generic binary characteristics.

2.2. Graph Theory

Maximizing only the intra-group efficiency for teams of equal size is equivalent to the NP-hard *balanced k-clique partitioning problem* (BCPP) (Bhasker and Samad 1991), which has been used, for example, for sports tournament scheduling and league realignment (Recalde et al. 2016). This stream of literature focuses only the efficiency of the solution, with no consideration of stability.

2.3. Matching

With preferences submitted by self-interested agents, the current work is related to an extensive literature on \matching markets" which has become an active area of research with wide-spread applications. This work builds on the classical *stable marriage problem* of Gale and Shapley (1962) (see also Gusfield and Irving 1989, Irving 1985, Iwama and Miyazaki 2008) in which disjoint sets of men and women each rank members of the opposite gender (possibly ranking \being alone" above some potential mates). A *stable* matching is often desired, in which no unmatched couple prefers each other over their current matching. Applications of bipartite marriage-type matching have flourished in recent years, with prominent successes in the National Residency Matching Program (Roth 1996), school-choice programs (Pais and Pint´er 2008), and kidney exchange (Roth et al. 2004), and even recognition of the discipline with the 2012 Nobel Prize in Economic Sciences. The tension between efficiency and stability (studied here) has been present in this stream of research, particularly in discussions of school-choice program implementations (Erdil and Ergin 2008).

2.3.1 Roommate Problem Variations

As mentioned in Section 1, when the market does not consist of two disjoint classes of agents (e.g., men and women, workers and jobs, students and schools, etc.) but is instead drawn from a single pool of agents, a stable solution may not exist, even with “teams" of size two and no further constraints, i.e., the *roommate problem* (Irving 1985). Indeed, team formation has already been discussed as a natural generalization of the roommate problem (Bir´o et al. 2016), but few have attempted to tackle all computational difficulties directly.

While a particular instance of the (s = 2) roommate problem may not admit a stable solution, a polynomial-time algorithm (Irving 1994) can check for the existence of a stable solution and find one, if it exists. Prosser (2014) present an alternative constraint-programming algorithm for this problem with partially-defined preference lists. Partially-stable outcomes in *unsolvable* roommate problems (problems without stable matchings) have been defined; for example, an *almost stable matching* consists of a Pareto-optimal matching with a minimum number of blocking pairs; a *maximum-internally stable matching* is a solution with a maximal set of stable pairs; and a *maximum irreversible matching* maximizes the *number* of stable pairs (Abraham et al. 2006, Tan 1990, Bir´o et al. 2016). Also, Van der Linden et al. (2016) recently explored an algorithm for the roommate problem in which people who are mutual favorites, i.e., *soulmates*, are matched first, with the process then iterated. To the best of knowledge of the authors, no formal study of determining whether or not a stable assignment exists (and producing one if it does exist) for s > 2 has previously been studied. Other generalizations of stable roommate problems include those by Cechl´arov´a and Fleiner (2005), where an agent may participate in more than one 2-person relationship. Recently, Wolfson and Lin (2017) study ride-sharing as an application of the roommate problem, proposing a heuristic algorithm for both efficiency and stability, but again studying only the special case of our current setting under s = 2.

It is worth noting here that much of the roommate and matching literature focuses on ordinal-preference elicitation (i.e., submission of ranked lists) while we use cardinal-preference elicitation (i.e., submission of numerical scores) for practical reasons. Because every cardinal submission can be transformed uniquely to a weak ordinal preference, we ignore this distinction for the remainder of the paper.

2.3.2 Hedonic Games

This paper focuses on team formation with a restriction to *equal-sized* teams, unlike *hedonic games* (Aziz et al. 2011). Negative preferences in hedonic games can result in agents not matched to teams, where here an equal-size constraint prohibits this. We thus normalize preferences to be nonnegative.

Also of note is that a variety of stability concepts have been studied for hedonic games, including *Nash stability*, *individual stability*, and *contractual individual stability*, respectively considering the benefit of leaving a current team and joining another team for each person of the market, for each person and the team receiving a defector, and for each person, the receiving team, and the team that was left (Aziz et al. 2011).

Research has been done on the complexity of verification, existence, and calculation of solutions satisfying these notions of stability in hedonic games (Sung and Dimitrov 2010), but the emphasis on individual deviations does not shed much light on the difficult computational problems explored here. The equal-size constraint forces us to consider group deviation directly. Because one individual cannot defect and join another team without implications for other teams, it is important to consider improvements resulting from several trades taking place at once. Our notion of *team uplift* considers the whole group of agents' preferences in forming a hypothetical alternative team. We consider individual deviation only in the appendix, showing that it seems to be less interesting with equal-size teams.

2.4 Team Formation: Benchmarks from the Literature

The most closely related work to our own is that of Wright and Vorobeychik (2015) who also explore mechanisms for team formation computationally. In the Harvard Business School Draft (**DRAFT**), agents are randomly ordered with the first m agents selected as captains. Over m - 1 rounds, each captain in turn selects her most preferred unassigned member to join her team, with the order reversing in even and odd numbered rounds. In the One-Player-One-Pick (**OPOP**) mechanism, all agents are randomly ordered with the first m captains starting a team with only one pick from among the remaining non-captains. Then, each remaining non-captain in the ordering chooses her favorite team based on expected utility. If her team has another available spot, she also chooses the next person to join her team. Wright and Vorobeychik (2015) provide evidence in favor of these two mechanisms, making them the best benchmarks available for direct comparison in our computational experiments.

3. The Team Formation Problem

Let $N = \{1, \dots, n\}$ be a set of n agents with $n = m \cdot s$ and $n, m, s \in Z^+$. For each $i \neq j \in N$, let $u_{i,j} \in Z^+$ represent the *pairwise utility* of i for being teamed with j, normalized to be non-negative and integer by affine transformation, with all $u_{i,j} = 0$. A feasible *team formation* is a partition of N into m teams $M = \{1, \dots, m\}$ denoted by $t: N \to M$. Hence $t(i)$ is the team agent i is assigned to, and the *equal-size teams* restriction requires $|\{i \in N | t(i) = k\}| = s$ for all $k \in M$. We will use T to denote the set of legal team formations, with the equal-size teams restriction assumed throughout. Furthermore, let $c(t, i) := \{j : t(j) = t(i)\}$ be the set of agents assigned to the same team as i in t.

It will be convenient to search over size-s subsets of N without specifying an entire team formation. Thus, let $C := \binom{N}{s}$ be the family of all subsets of N (called *coalitions*) of size s. For any i, let $C(i) \subseteq C$ be those coalitions c with $i \in c$. For any $c \in C(i)$, let $u(c;\ i) := \sum_{j \in c} u_{i,j}$ signify i's total individual utility as part of c. As additional shorthand, we have $u(t;\ i) := u(c(t, i), i)$ as i's utility in a team formation and $u(c) := \sum_{i \in c} \sum_{j \in c} u_{i,j}$ as the utility of an entire coalition c. We define each agent i's maximum realizable utility $U(i) := \max_{c \in C(i)} u(c, i)$, which can be found by simply taking the top $s - 1$ values of $u_{i,j}$ (until later when other constraints beyond equal-size are added).

As motivated above, we are interested in team formation with both total utility (efficiency) and stability as objectives. Thus, our formal optimization problem, the *team formation problem* (TFP) is formulated with both objectives, weighted by a scalar $\alpha \in [0,1]$. The first component (weighted by α)

seeks to maximize the sum of the individual utilities. The second component (weighted by 1 - α) seeks to minimize the maximum uplift r, defined for any fixed $t \in T$ by a maximum of

$$r(c,t) := \sum_{i \in c} \big(u(c,i) - u(t,i)\big),$$

over all coalitions $c \in C$ with $u(c,i) \geq u(t,i)$ for all $i \in c$. The TFP is thus modeled as:

$$\max_{t \in T} \alpha . \textstyle\sum_{i \in N} u(t,i) \; - \; (1-\alpha) \cdot r \qquad (TFP)$$

$$s.t. \quad r = \max_{c \in C} \{ r(c,t) : (i \in c) \rightarrow (u(c,i) \geq u(t,i)) \}$$

The upper-level (leader) optimization model selects the partition and the lower-level (follower) optimization model calculates the maximum uplift coalition for the partition identified in the leader model. The condition in the follower problem enforces that each individual in a maximum uplift coalition is not worse off.

Note that if α = 1 we arrive at the BCPP, and if α = 0 then stability (as measured by maximum uplift) is the only measure of interest. In the latter case, if the optimal solution is 0, then we have a fully stable solution; if positive, we have a solution that minimizes the maximum uplift.

Optimization model (TFP) can be cast as a bi-level binary optimization problem by introducing a binary variable $x_{i,k}$ indicating if each agent i is placed in team k. We also associate, in the follower, variables y_i, to indicate if person i is selected in the maximum uplift coalition:

$$\max \quad \alpha \cdot \textstyle\sum_{i \in N} \sum_{j \in N} \sum_{k \in M} u_{i,j} x_{i,j} . x_{j,k} - (1-\alpha) \cdot r \qquad (BL)$$

$$s.t. \quad \textstyle\sum_{i \in N} x_{i,k} = s, \qquad \forall k \in M$$

$$\textstyle\sum_{i \in N} x_{i,k} = 1 \qquad \forall i \in N$$

$$x_{i,j} \in \{0,1\} \qquad \forall i \in N, \forall k \in M$$

$$r = \max\{ \textstyle\sum_{i \in N} \sum_{j \in N} u_{i,j} y_i y_j - \sum_{i \in N} \sum_{j \in N} \sum_{k \in M} y_i u_{i,j} x_{i,k} x_{j,k}$$

$$s.t. \quad 0 \leq (1 - y_i) \cdot \boldsymbol{U}(i) + \sum_{j \in N} u_{i,j}\, y_i y_j - \sum_{j \in N} \sum_{k \in M} y_i u_{i,j} x_{i,k} \qquad \forall i \in N$$

$$\sum_{i \in N} y_i = s$$

$$y_i \in \{0,1\} \qquad \forall i \in N\}$$

Proposition 1. *Model (BL) is a valid formulation for the TFP.*

Proof. Leader problem constraints ensure a team formation, with total utility weighted by α in the objective. We need only show that the follower problem identifies the maximum uplift group.

Fix a solution to the leader problem x'. Let y' be a feasible solution to the follower. By the constraint $\sum_{i \in N} y_i = s$, a properly sized coalition is identified. The first term in the follower objective function, $\sum_{i \in N} \sum_{j \in N} u_{i,j} y_i y_j$ is the total coalitional utility over every i with y'_i= 1. The second term evaluates to $\sum_{i \in N} \sum_{j \in N} \sum_{k \in M} y'_i u_{i,j} x'_{i,k} x'_{j,k}$Therefore, for a fixed i, $y'_i u_{i,j} x'_{i,k} x'_{j,k}$ will evaluate to 0 if $y'_i = 0$, and to $\sum_{j \in N} \sum_{k \in M} u_{i,j} x'_{i,k} x'_{j,k}$ if $y'_i = 1$. In the latter case, this term is the utility of *i* according to x' in the leader problem. Therefore, the (1 - α) objective term evaluates to the total uplift of the coalition defined y'_i above the *x'* value.

Finally, the constraint $0 \leq (1 - y_i) \cdot U(i) + \sum_{j \in N} u_{i,j}\, y_i y_j - \sum_{j \in N} \sum_{k \in M} y_i u_{i,j} x_{i,k}$ becomes trivially satisfied if $y'_i = 0$ (by the maximality of $U(i)$). If $y'_i = 1$, the constraint enforces that each individual chosen must have an increase in utility to want to join an uplift coalition, completing the proof. □

4. Single-level Reformulations

In order to design an efficient algorithm for solving (BL) (which is bi-level, quadratically constrained and has a quadratic objective function), we transform it into a single-level linear binary optimization model with the addition of two new sets of binary variables. For every pair $i, j \in N$, let $w_{i,j}$ indicate if *i* and *j* are assigned to the same team (if $i = j$ let $w_{i,j}$ = 0, or simply ignore it). Next, for every *i* and value $v \in \{0,1, \dots, U(i)\}$, introduce variables $z_{i,v}$ indicating if *i* has individual utility *v* in the team assigned to *i*:

$$\max \quad \alpha \cdot \sum_{i\in N}\sum_{j\in N} u_{i,j} w_{i,j} - (1-\alpha)\cdot r \qquad (SL)$$

$$s.t. \quad \sum_{i\in N} x_{i,k} = s, \qquad \forall k \in M$$

$$\sum_{i\in N} x_{i,k} = 1 \qquad \forall i \in N$$

$$w_{i,j} \geq x_{i,k} + x_{j,k} - 1 \qquad \forall i \in N, \forall j \in N, \forall k \in M$$

$$w_{i,j} + x_{i,k} \leq x_{j,k} + 1 \qquad \forall i \in N, \forall j \in N, \forall k \in M$$

$$\sum_{v=0}^{U(i)} z_{i,v} = 1 \qquad \forall i \in N$$

$$\sum_{v=0}^{U(i)} v.z_{i,v} = \sum_{j\in N} u_{i,j} w_{i,j} \qquad \forall i \in N$$

$$r \geq u(c) - \sum_{i\in c}\sum_{v=0}^{U(i)} v.z_{i,v} - \sum_{i\in c}\sum_{v=u(c,i)+1}^{U(i)} u(c).z_{i,v} \qquad \forall c \in C$$

$$x_{i,j} \in \{0,1\} \qquad \forall i \in N, \forall k \in M$$

$$w_{i,j} \in \{0,1\} \qquad \forall i \in N, \forall j \in N,\ i \neq j$$

$$z_{i,v} \in \{0,1\} \qquad \forall i \in N, \forall v \in \{0,1,..,U(i)\}$$

Theorem 1. *Model (SL) is a valid formulation for the TFP.*

Proof. The first four sets of constraints generate a partition and link the *x* variables with the *w* variables. Note that the second constraint set is necessary to link the *x* and *w* variables, forcing there to be at least one *k* for which $w_{i,j}$ is constrained to zero when *i* and *j*'s teams differ.

The constraints $\sum_{v=0}^{U(i)} z_{i,v} = 1$ and $\sum_{v=0}^{U(i)} v.z_{i,v} = \sum_{j\in N} u_{i,j} w_{i,j}$ define the *z* variables, which select exactly one value from the domain of the agent's utility function and match it to the *w* variable's generation of utility as the sum of pairwise actualizations.

The final constraint set (aside from binary constraints) links the uplift variable *r* to the uplift of each coalition. In particular, for every coalition *c*, the first term on the right-hand side is the total utility for *c*.

The second term calculates, for each *i* in the coalition, the individual utility realized in the team formation obtained endogenously (by *x* and hence *w* and *z* variables). The third term ensures that the constraint does not restrict the value of *r* unless it is an *individually rational* uplift group for all $i \in c$. That is, if $z_{i,v} = 1$ for a value $v > u(c, i)$, then *i* would get more utility from the endogenous team formation than from the potential coalition c, and so would not join c. An active $z_{i,v}$ variable in the third term causes the entire right-hand side to be non-positive, hence trivially satisfied; if even one agent would not join c it is not considered as a stability threat. Note that these constraints enforce $r \geq 0$; for any group c in the solution defined by *x*, the right-hand side evaluates to 0. With the coefficient of *r* in the objective function non-positive, *r* will take the maximum value of the right-hand sides over this constraint set. Any tight constraint from this set corresponds to a maximum uplift coalition in the solution defined by *x*. □

An alternative, exponentially sized model can be formulated, which provides a tighter linear relaxation, at the expense of model size. This is done by creating a binary variable t_c for every $c \in C$, indicating whether or not this coalition is part of the team formation.

$$\max \quad \alpha \cdot \textstyle\sum_{c \in C} u(c) t_c - (1-\alpha) \cdot r \qquad (EXP)$$

$$s.t. \quad \textstyle\sum_{c \in C} t_c = m, \qquad \forall k \in M$$

$$\textstyle\sum_{c \in C(i)} t_c = 1 \qquad \forall i \in N$$

$$w_{i,j} = \textstyle\sum_{c \in C(i) \cap C(j)} t_c \qquad \forall i \in N, \forall j \in N, i \neq j$$

$$\textstyle\sum_{v=0}^{U(i)} z_{i,v} = 1 \qquad \forall i \in N$$

$$\textstyle\sum_{v=0}^{U(i)} v.z_{i,v} = \textstyle\sum_{j \in N} u_{i,j} w_{i,j} \qquad \forall i \in N$$

$$r \geq u(c) - \textstyle\sum_{i \in c} \sum_{v=0}^{U(i)} v.z_{i,v} - \sum_{i \in c} \sum_{v=u(c,i)+1}^{U(i)} u(c).z_{i,v} \qquad \forall c \in C$$

$$t_c \in \{0,1\} \qquad \forall c \in C$$

$w_{i,j} \in \{0,1\}$ $\forall i \in N, \forall j \in N, i \neq j$

$z_{i,v} \in \{0,1\}$ $\forall i \in N, \forall v \in \{0,1,..,U(i)\}$

This model replaces the *x* variables with *t* variables to define the team formation, and the *x* and *w* linking constraints with $w_{i,j} = \sum_{c \in C(i) \cap C(j)} t_c$ linking *t* variables with *w* variables.

4.1 Characteristic Constraints

As noted in section 2 diversity of skill-sets or types is a primary concern when establishing teams. For example, a company may want teams with individuals from diverse functional areas, with varying *Myers-Briggs Type Indicator* scores (The Myers & Briggs Foundation 2018), or with different expertise. With student teams, diversity with respect to gender, skills, or roles might be desired.

In general we consider constraints of the form, \each team must have a specified minimum (or maximum) number of agents with characteristic q." Formally, suppose that there are a set of characteristics Q that each individual will either possess of not. This is indicated by the binary parameter $\delta_{i,q}$ equal to 1 if and only if agent *i* has characteristic q. For each characteristic $q \in Q$, there is a specified $quota_q$ and cap_q, with $0 \leq quota_q \leq cap_q \leq s$, establishing bounds on the number of individuals possessing each characteristic that must be represented in each group.

These bounds are easily appended to the models above. For model with *x* variables, namely models (BL) and (SL), add:

$$quota_q \leq \sum_{i \in N} \delta_{i,q}, x_{i,k} \leq cap_q \qquad \forall q \in Q, \forall k \in M$$

to enforce this condition. For (SL) and (EXP), with variables t_c, we simply refine *C* to contain only those groups *c* for which these conditions hold. Call this set C^Q.

5. Optimization Algorithms

The models (SL) and (EXP) are computationally challenging in practice. Both contain, at a minimum, a pseudo-polynomial number of variables and an exponential number of constraints. This requires the design of algorithms capable of scaling to instances of practical size. We make use of branch-and bound search, as is typically employed for binary optimization problems, with added routines for handling the exponentially sized portions of the model.

This section provides a description of two optimization algorithms for TFP, one designed to solve model (SL) and one for model (EXP). To begin, we describe a class of optimization problems that will appear as subproblems throughout these two primary algorithms. The *cardinality-constrained binary quadratic programming problem* (CCBQP) is specified by an asymmetric, not necessarily positive semi-definite, $\nu \times \nu$ matrix Q, and a value K:

$$\max \chi^T Q \chi \qquad (CCBQP)$$

$$\begin{aligned} s.t. \quad & \textstyle\sum_{i \in N} \chi_i = K \\ & A\chi \geq b \\ & \chi_i \in \{0,1\} \qquad \forall i \in \{1, \dots, \nu\} \end{aligned}$$

where the additional linear constraint set $A\chi \geq b$ may be vacuous. Recent literature has investigated effective computational models for solving various instance types of CCBQP problems. In particular, model 2 of Lima and Grossmann (2017) proved most efficient in our experimental results (using commercial IP solvers) over a wide-range of instances, and so our presented results employ the following reformulation throughout. Introduce binary variables $\psi_{i,j}$ for every pair of indices $i, j \in \{1, \dots, \nu\}$ with $i < j$, and reformulate (CCBQP) as follows:

$$\max \textstyle\sum_{i=1}^{\nu} Q_{i,i}\, \chi_i + \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} (Q_{i,j} + Q_{j,i})\, \psi_{i,j} \qquad \text{(CCBQP*)}$$

$s.t.\ \ \sum_{i=1}^{v} \chi_i = K$

$\psi_{i,j} \geq \ \chi_i + \chi_j - 1$

$\psi_{i,j} \leq \ \chi_i$

$\psi_{i,j} \leq \ \chi_j$

$\sum_{i=1}^{j-1} \psi_{i,j} + \sum_{i=j+1}^{v} \psi_{i,j} \ = (K-1)\ \chi_i \qquad \forall j \in \{1, \dots, v\}$

$A\chi \geq b$

$\chi_i \in \{0,1\} \qquad \forall i \in \{1, \dots, v\}$

We now describe three proposed algorithms for the TFP, each a branch- and-bound algorithm utilizing a model formulated in Section 4.

5.1.Branch and cut (BC)

BC is a branch-and-cut algorithm for solving model (SL). All computational models for the TFP presented in this paper contain the family of uplift- defining constraints:

$$r \geq u(c) - \sum_{i \in c} \sum_{v=0}^{U(i)} v.z_{i,v} - \sum_{i \in c} \sum_{v=u(c,i)+1}^{U(i)} u(c).z_{i,v} \qquad \forall c \in C \qquad (UP(C))$$

As opposed to adding all $UP(c)$ constraints at once, we propose a *branch- and-cut* approach for finding constraints that might impact the optimal solution. Namely, at each integer-search-tree node, an optimization problem is solved to identify if there exists any $c \in C$ for which the *c*-indexed constraint $UP(c)$ is violated. If such a *violated constraint* exists, it is added to the model and the branch-and-bound search continues.

Proposition 2. *At any integer search-tree node, for either (SL) or (EXP), let r_j be the value of r and, $\forall i \in N$, let v_{ij} be the unique second index for which variable $z_{i,v}$ is1. Let v_i' be the optimal solution to the following problem, with optimal objective value r^*:*

$$\max \sum_{i\in N}\sum_{j\in N} u_{i,j}\ \chi_i - \sum_{i\in N} v'_i\ \chi_i \qquad (VC)$$

$$s.t.\ \sum_{i=1}^{n} \chi_i = s$$

$$\sum_{j\in N} u_{i,j}\ \chi_i \geq v'_i\ \chi_i \qquad \forall i \in N$$

$$\chi_i \in \{0,1\} \qquad \forall i \in N$$

There exists a violated $UP(c)$ constraint if and only if $r^* > r'$. Furthermore, if $r^* > r'$, and $c^* := \{i : \chi_i^* = 1\}$, then $UP(c^*)$ is violate constraint.

The proof of this proposition is immediate the mathematical program (VC) finds the group c^* for which each individual has a non-decreasing uplift (enforced by the constraints $\sum_{j\in N} u_{ij}\,\chi_i \geq v'_i.\chi_i$ corresponding to the most-violated constraint. Model (VC) is a special case of the $CCBQP$ and so can be solved via model $(CCBQP^*)$. We note that if characteristic constraints are considered they can be directly added to the mathematical program in the statement of Proposition 2 by adding constraints:

$$quota_q \ \leq\ \sum_{i\in N} \delta_{i,q}\chi_i \leq\ cap_q \qquad \forall q \in Q$$

Equipped with Proposition 2 we can formally describe our first proposed algorithm BC, which solves model (SL) by branch-and-cut. Starting with none of the constraints $UP(c)$, a branch-and-bound search solves model (SL). At any integer-search-tree node, the optimization model in Proposition 2 is solved to find if there exists a violated constraint. If one exists, the constraint $UP(c^*)$ is added to the model and the search continues. Otherwise, the solution identified at the node is feasible, and a potentially improving solution.

5.2. Branch-cut-and-price (BCP)

BCP is a branch-cut-and-price algorithm for solving (EXP). One simple method, which we denote by EXP, is enumerating all of *C* (or C^Q if *caps* or *quotas* are in use) and directly solving (EXP) using a state-of-the-art integer programming solver. EXP has the obvious shortcoming of scalability—$|C|$ grows exponentially with the number of people, hence limiting the use in practical application.

To address this shortcoming, we describe a branch-cut-and-price algorithm, BCP, that generates variables and constraints dynamically, as needed. Note that not only does model (EXP) contain an exponential number of variables and an exponential number of constraints, but there is a family of constraints that contain exponentially many variables, requiring particular care in implementation. We will interchangeably refer to the variables t_c as *variables or columns*.

BCP is initialized by finding any set of coalitions $C^0 \subseteq C$ representing a team formation. For any current active $C' \subseteq C$, starting with C^0, define the *restricted master problem*, *RMP(C′)*, as exactly formulation (EXP) with C replaced by C', implying a restricted set of both t_c columns and $UP(c)$ constraints. Let $o^*(C')$ denote the optimal value of $RMP(C')$, with $r^*(C')$, $t_c^*(C')$, and $w_{i,j}^*(C')$ as the associated variables values at the particular optimal solution. To isolate the efficiency component of the objective, let $u^*(C') := \frac{o^*(C')+(1-\alpha)r^*(C')}{\alpha}$. Note that with $C' = C$, all these values correspond to the true optimal solution (of the unrestricted problem), in which case we may drop the argument. For the linear relaxation of $RMP(C')$ (replacing each $\in \{0, 1\}$ with $\in [0, 1]$), we denote the optimal value and solution by adding carets (or hats)—for example, $\hat{o}^*(C')$ refers to the optimal value of the linear relaxation of $RMP(C')$.

With C^0 defined as a feasible team formation, $RMP(C^0)$ is feasible. Yet $o^*(c^0)$ may not be a lower bound on o^*, nor will $\hat{o}^*(C^0)$ necessarily be an upper bound on o. For the former, there may be additional constraints, related to $c \notin C^0$, that restrict r to take a value higher than it does at the solution. For the latter, there may be variables that needed to be added that are members of an improving solution.

Fix C' and consider the linear relaxation of $RMP(C')$. Let μ, σ_i, *and* $\kappa_{i,j}$ be the dual multipliers of the first three listed constraints at the optimal value, defined on the appropriate indices. Theorem 2 establishes a condition for which one can assert that is an upper bound on o^*, enabling an exact branch-cut and-price algorithm to be designed.

Theorem 2. Let rc^*be the optimal value to the following problem:

$$\max \sum_{i\in N}\sum_{j\in N\ j\neq i}(u_{i,j}-\kappa_{i,j})\chi_i\chi_j - \sum_{i\in N}\sigma_i\chi_i - \mu \qquad (RC)$$

$$s.t. \sum_{i=1}^{n}\chi_i = s$$

$$\chi_i \in \{0,1\} \qquad \forall i \in \{1,\ldots,n\}$$

If $rc^* \leq 0$ and the optimal value to (VC) is less than or equal to $r^*(C')$, then $\hat{o}^*(C')$ is an upper bound of o^*.

Proof. We need only show that under the conditions of the theorem there is no set of groups $\tilde{C} \supset C'$ for which $\hat{o}(\tilde{C}) > \hat{o}(C')$. Let P^* be the problem formed by adding to $RMP(C')$ any missing $UP(c)$ constraints $\forall c \in C$ but keeping the index set $C^{'}$ for t_c columns. Given this set of columns defined by $C^{'}$, since the optimal value to $RMP(C')$ is less than or equal to $r^*(C')$, we know that the optimal value of the LP relaxation of $RMP(C')$ is equivalent to the optimal value of the LP relaxation of P^*. We now show that, under the conditions of the theorem, for any coalition $\tilde{c} \notin C'$, the reduced cost of variable $t_{\tilde{c}}$, if it were to be added to the variables in P, is less than or equal to 0. This implies that the LP relaxation of P^* is equal to $\hat{o}^*(C')$, which in turn implies $\hat{o}^*(C') = \hat{o}(C)$.

Fix $\tilde{c} \notin C'$. The reduced cost of $t_{\tilde{c}}$ is

$$u(\tilde{c}) - \mu - \sum_{i\in\tilde{c}}\sigma_i - \sum_{i\in\tilde{c}}\sum_{j\in\tilde{c}:j\neq i}\kappa_{i,j} = \sum_{i\in\tilde{c}}\sum_{j\in\tilde{c}:j\neq i}\left(u_{i,j}-\kappa_{i,j}\right) - \sum_{i\in\tilde{c}}\sigma_i - \mu$$

By the conditions of the theorem, $\sum_{i\in\tilde{c}}\sum_{j\in\tilde{c}:j\neq i}\left(u_{i,j}-\kappa_{i,j}\right) - \sum_{i\in\tilde{c}}\sigma_i - \mu \leq 0$. Since $\tilde{c}$ was arbitrarily chosen, this concludes the proof.□

Note that RC is another instance of the $CCBQ$.

Equipped with Theorem 2, BCP proceeds as follows. Starting from an appropriate C^0, we begin by iterating between finding improving columns, and finding violating constraints. In particular, starting with $C' = C^0$ (RC) (the pricing problem) is iteratively solved, and whenever an improving column is found (i.e., the optimal value is greater than 0), it is added to C'. When the optimal value drops to 0, model (VC) is solved. If a violating constraint is found, the coalition corresponding to that constraint is added to C', and, again, improving columns are identified. If, however, there are no violating constraints, a global bound on the optimal value is found. One can then solve $RMP(C')$ with integrality constraints to arrive at a feasible, potentially improving primal solution. This is done at every search-tree node, as $RMP(C')$ can typically be solved effectively.

A branch-and-bound search is used to find the globally optimal solution. In particular, after solving the root node of the search, by the procedure described above, a variable is chosen to branch on. Two other nodes are created, one in which the chosen variable is forced to 0 and the other forced to 1. The procedure continues until all nodes are pruned, whereupon the best found primal solution is the globally optimal solution to the original TFP instance. In the case where characteristic considerations are incorporated, one need only add constraints to the pricing problem limiting the choice of additional columns.

More formally, a branch-and-bound search is implemented as follows. Each search-tree node is specified by two coalitional sets: C^{in} and C^{ex}. Coalitions $c \in C^{in}$ are required to be a part of the solution and $c \in C^{ex}$ are forced to be excluded from the solution, and not be generated in subsequent pricing problems. The search-tree nodes are stored in a queue L. Initially, a root node ρ contains $C^{in} = C^{ex} = \emptyset$, a relaxation value o is assigned to ρ, and $L := \{\rho\}$.

Once processing at a search-tree node concludes, a new search node ρ is selected from L. The LP relaxation of (EXP) is found by solving $RMP(C')$ with added constraints enforcing each $c \in C^{in}$ to be in the solution (i.e., $t_c = 1$) and that each $c \in C^{ex}$ is excluded (i.e., $t_c = 0$). This is solved by iteratively finding new columns to add to $C^{'}$ via (RC) (adding constraints to exclude any $c \in C^{in} \cup C^{ex}$ by adding the constraint $\sum_{i \in c} \chi_i - \sum_{i \notin c} \chi_i \leq s - 1$, and then finding the most violated constraint by (VC). By Theorem 2, if no improving column and no violated constraint is found, the LP relaxation is solved and a valid upper bound for the search node $o'^{'}$ is found.

If o' is less than or equal to the value of the best known feasible solution, the node is pruned and search continues. Otherwise, the master problem $RMP(C')$ is solved with the integrality constraints included (along with constraint enforcing group include/exclusion in C^{in} and C^{ex}). If this solution is the best known, it is recorded.

If o' is still strictly larger than the value of the objective value of the best known solution, two descendant nodes are created by selecting a coalition $c' \in C'$ for which $0 < t_{c'} < 1$. In one node c' is added to C^{in} and in the other it is added to C^{ex}.

If $|L| = 0$, the best known solution must be optimal.

5.3. Implementation details

We implemented techniques BC, EXP, and BCP on an Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz and 8 GB RAM, written in C++ and compiled with GCC 4.8.4. IP models were solved with Gurobi 7.5.1. Unless otherwise noted, the experiments below all use an 1800 second time limit. All instances used in the experimental evaluation and MATLAB instance generators are available upon request. Further details are as follows.

BC: solved by GUROBI with cuts (VC) identified via a callback. Through preliminary experimentation, the following settings for GUROBI were found most effective: PreCrush = 1; DualReductions = 0; LazyConstraints = 1. Cuts $UP(c)$ are found by solving (VC) with the reformulation into model (CCBQP*).

EXP: solved by GUROBI with default setting via enumerating all of C. BCP: solved by starting with an arbitrary solution (found by taking the first s and putting them in one group, then the next s, etc.), and adding these groups to C^0. We then select 100 other random groups to add to C^0 as follows. For each of these teams, we first select a random person. Then, Then, having selected a group g' of size $s' < s$, each unselected person $j' \notin g'$ is assigned probability $P_{g'}(j') = \frac{\sum_{i \in g'} u_{i,j'}}{\sum_{i \in g'} \sum_{j \notin g'} u_{i,j}}$ and is chosen to be the next person added to g' with probability $p(j')$. This process continues until a group of size s is found and then that group is added to C^0. This process is repeated 100 times.

In the branch-and-bound search, the following algorithmic specifications are set. Each pricing problem is solved by GUROBI with Cuts = 0 (this was found most effective in preliminary computational results) using reformulation CCBQP*. The nodes L are stored in a priority queue and the node with the largest LP relaxation of the parent that created the node is selected. The group with the most fractional $t_{g'}$ is chosen to branch on.

6. Experimental Evaluation of EXP, BC, and BCP

Our results can be summarized as follows, with details provided in each corresponding subsection:

• **BCP** tends to dominate **EXP** and **BC** on the number of instances solved and optimality gap, though **BC** can at times provide an optimality bound when the others cannot. (§6.2)

• Under **BCP**, larger u^{max} values tend to result in more easily solved instances, despite a larger pseudo-polynomial formulation. (§6.3)

• The boundary values $\alpha \in \{0, 1\}$ result in inferior solutions and should be avoided in favor of, for example, $\alpha \in \{0.01, 0.99\}$. A small con- sideration of efficiency finds more/better stable solutions in the same amount of time, or conversely, a small consideration of stability finds more stable solutions of equal efficiency with only a small amount of additional computation. (§6.4)

•The type of characteristic constraints generated here have a significant impact on stability but an insignificant effect on total utility. If such constraints are desired, Pareto improvements may often be available. (§6.5)

Based on these results, the comparison of our own techniques to the benchmarks from the literature (in §7) restricts attention to **BCP** with $\alpha \in \{0.01, 0.5, 0.99\}$.

6.1Instance Generation

To explore the multidimensional parameter space, we focused on markets of size (m, s) ∈ {(2, 4), (2, 12), (3, 8), (4, 2), (4, 4), (4, 6), (4, 8), (6, 4), (8, 4)}, including therefore instances with $n \in \{8, 16, 24, 32\}$. For each (m, s)-pair we generate 20 instances (five for each of the four preference generation models described below) using $u^{max} = 25$. Further, for (m, s) = (6, 4) we also generated twenty instances each for $u^{max} \in \{5, 100\}$, for use in §6.3, a study of the effect of the u parameter. Finally, for each these eleven market parameter settings (nine with $u^{max} = 25$ and one each with $u^{max} = 5, 100$) we also generate a parallel set of instances with three personal characteristics (i.e., with $|Q| = 3$). This doubles the number of instances, resulting in 440 simulated markets, 220 each for $|Q| \in \{0, 3\}$.

To ensure feasibility when supplementing an instance with characteristic constraints, δ_i, q values are determined by solving the following optimization problem:

$$\max \sum_{i \in N} \sum_{q \in Q} w_{i,q} \cdot \delta_{i,q} \qquad (FB)$$

$$s.t. \quad quota_q \leq \sum_{i \in N | t(i)=k} \delta_{i,q} \leq cap_q \qquad \forall q \in Q, \forall k \in M$$

$$\delta_{i,q} \in \{0,1\}$$

For any values of $w_{i,q}$, this model produces $\delta_{i,q}$ values so that at least one valid team exists. For each instance, values $w_{i,q} \sim U[-1, 1]$ are drawn independently $\forall i \in N$ and $\forall q \in Q$. This provides a random assignment of all three personal binary characteristics such that quotas and caps as defined in §4.1 are always feasible.

Proceeding, each market is defined by $u_{i,j}$ values, which can be generated randomly according to a preference generation model. Here we employ two models from the literature (**G1** and **G3**) and offer two new variants of our own (**G2** and **G4**).

Monotone Common-Value (G1): $u_{i,j} = j + \varepsilon_{i,j}$, where $\varepsilon_{i,j} \sim N\left(0, \frac{n}{5}\right)$. (Note that $\varepsilon_{i,j}$ is re-sampled until $u_{i,j}$ is a positive number.) This recreates the generation procedure of Othman et al. (2010), with a common-value component j (equal to the agent index) and a normally distributed private deviation from the common baseline opinion of an agent's value. (This may simulate, for example, a sports draft where the relative order of players does not vary much from one selecting team to the next, resulting in highly correlated preferences.)

Clustered Common-Value (G2): $u_{i,j} = l_j + \varepsilon_{i,j}$, where the common value l_j has four levels, $l_j \in \{0, \frac{n}{4}, \frac{n}{2}, \frac{3n}{4}\}$ for four equal segments of the market. The resulting preferences are similar to G1, but with a heavier reliance on the private values to distinguish individuals. This models a market with less extreme public agreement on the value of agents than G1.

Uniform independent preferences (G3): $u_{ij} \sim U\{0, 100\}$, drawn independently from a discrete uniform distribution. For each agent, $n - 1$ numbers are chosen and sorted. Then, the differences between consecutive draws, in sorted order, are used to describe the utility of person *i* for other agents in the market (Wright and Vorobeychik 2015).

Affinity Preferences (G4): We designed this set of instances to model settings in which the characteristics of agents affect their utilities. First, characteristic sets are randomly generated based on model (FB). We assume that the first two characteristics represent traits for which like agents tend to have a natural affinity (e.g., age or gender) while the third characteristic is viewed favorably by all (e.g., higher skill level). We therefore generate $u_{ij} = \beta_{1,i} . \gamma_{i,j,1} + \beta_{2,i} . \gamma_{i,j,2} + \beta_{3,i} . \delta_{j3} + \varepsilon_{i,j}$ where binary $\gamma_{i,j,q} = 1$ and if and only if $\delta_{i,q} = \delta_{jq}$ for $q \in \{1, 2\}$. Further, independently drawn personal parameters $\beta_{1,i}$, $\beta_{2,i} \sim U[-1, 2], \beta_{3,i} \sim U[0, 2]$, and $\varepsilon_{i,j} \sim N(0, 1)$ reflecting that affinity effects are twice as likely to attract likes values as repel them (and to varying degrees), and that characteristic 3 is always viewed favorably but to varying degrees.

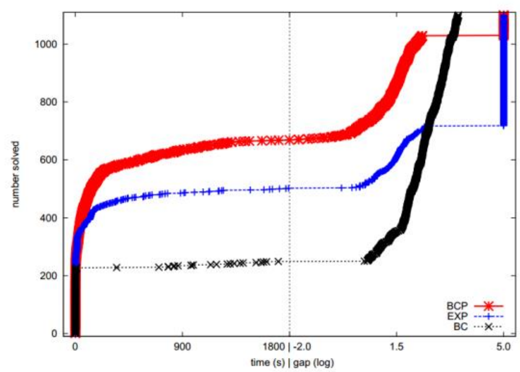


Figure 1: Cumulative distribution plot of performance, comparing the three algorithms developed in this paper.

In all four data-generation schemes, an $n \times n$ utility matrix U is generated as above and then *ipsative* scores are generated by normalizing with respect to the mean and standard deviation of the utility vector of agent i, then rescaled and rounded to make utilities positive, integer, and no more than u^{max} (Baron 1996).

6.2. Optimality, Gap, and Computation Time Analysis

We compare the relative efficiency of **EXP, BC**, and **BCP**, running each of the 220 unconstrained (i.e., |*Q*| = 0) market instances under $\alpha \in \{0, 0.01, 0.5, 0.99, 1\}$, resulting in 1100 runs of each algorithm. Detailed solution statistics appear in the Appendix, while Figure 1 depicts a cumulative distribution plot of performance. For each algorithm, the left half provides a plot with height equal to the cumulative number of instances solved at the time given on the horizontal axis. In the right half, the height of the plot corresponds to the number of remaining instances (unsolved at 1800s) with at most the log absolute gap (i.e., log(UB − LB)) shown on the horizontal axis. As a convention, we show an absolute gap of 100,000 (log of which is 5) for those

Table 1: Number of instances solved in 1800 seconds by **BCP** and **EXP**

	G1	**G2**	**G3**	**G4**
BCP	131	133	157	145
EXP	89	89	94	91

instances without a definite absolute gap (i.e., for **EXP** if memory limit is hit and for **BCP** if no upper bound is proven, meaning the root node is unresolved).

Figure 1 provides clear evidence of the superiority of **BCP** over **EXP**. **BC** is the most robust, in that it can provide a gap for all instances tested, being an entirely memory-controlled procedure. However, both **BCP** and **EXP** solve many more instances and leave a much smaller relative gap.

In summary, aggregated over these 1100 runs, **BCP** identifies a strictly better solution (lower bound) than the other two algorithms in 420 instances, and proves a strictly tighter relaxation bound (upper bound) in 344 in- stances. In comparison, **BC / EXP** provide strictly best lower bound and upper bound, respectively, in 36 / 57 and 71 / 43 instances. For those in- stances solved to optimality by both **BC** and **EXP**, the solution times for **EXP** are far superior and so we use **EXP** for the remaining comparisons to **BCP**.

Figures 2 through 5 provide more detailed comparison of **BCP** and **EXP** through scatter plots, one for each of the four generation schemes. In particular, for each generation scheme, a scatter plot consisting of a point per run (275 runs) with coordinates given by solution time of **BCP** and **EXP** is depicted. The size of each point corresponds to the number of agents in the instance. The color (gradient, from red to blue) corresponds to the team size, s. The point style corresponds to the value of α, with axes depicted in logscale.

These figures show that for the instances mutually solved, the solutions times are in general comparable, but that there are many instances unsolved by **EXP** that are solved by BCP, for all generations schemes. This is more apparent for **G3** and **G4** than for **G1** and **G2**, where **BCP** is able to solve even more instances. Table 1 reports the number of instances solved by **BCP** and **EXP**, respectively. This table suggests that **EXP** is only slightly affected by the generation scheme, but that **BCP** is able to solve significantly more instances for **G3** and **G4**, which suggests that varying degrees of common-value dependency makes instances more challenging.

The scatter plots also clearly exhibit that **EXP** is only superior to **BCP** when the number of agents and the size of the groups are relatively small.

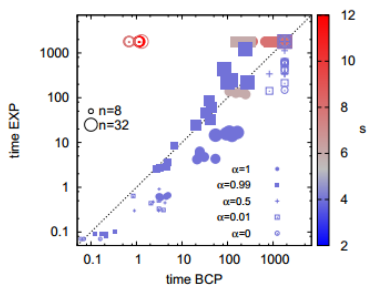


Figure 2: Runtime for **G1**

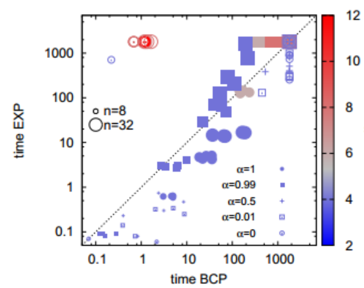


Figure 3: Runtime for **G2**

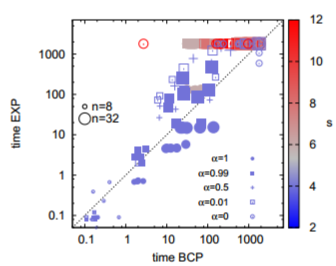


Figure 4: Runtime for G3

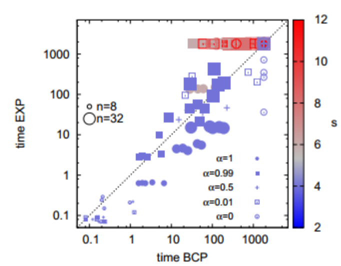


Figure 5: Runtime for G4

This coincides with expectations, since **EXP** is a precise model of the problem -however, as the problem size grows on any dimension, the application of **EXP** becomes prohibitive due to either memory restrictions or, even when memory limits are sufficient, resolution difficulty.

Based on the analysis in this subsection, we use the solutions obtained by **BCP** for the subsequent analysis. Note that even if an upper bound is not proven by **BCP**, a lower bound (i.e., high-quality feasible solution) can still be obtained by solving for the best solution using the columns generated. This contrasts with **EXP**, where, if the memory limit hits, no feasible solution will be available.

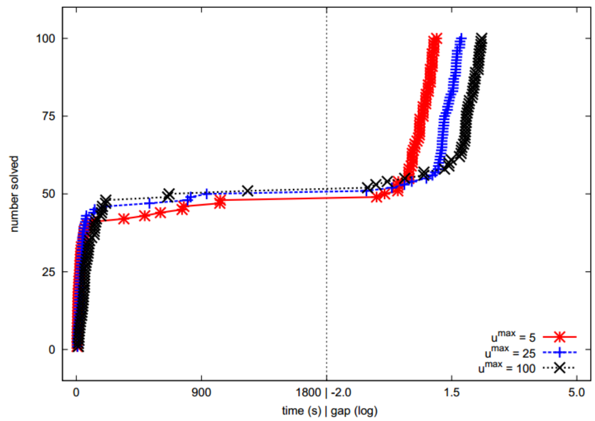


Figure 6: Cumulative distribution plot of performance, comparing **BCP** for varying u^{max} with m = 6 and s=4.

6.3. The effect of u^{max}

We next investigate the sensitivity of **BCP** to u^{max}. Figure 6 provides a cumulative distribution plot of performance for $u^{max} \in \{5, 25, 100\}$ on all instances generated with m = 6 and s = 4. (Four preference models, each with five instances and five values of α, result in 100 runs for each of three u^{max} values.) The figure exhibits an interesting relationship. Parameter u^{max} sets the number of possible values that an agent can specify for the other individuals in the market. Given more preference resolution/detail, it is slightly easier to find the optimal solutions. Additionally, the absolute gap increases for those instances that are unsolved. This can be attributed to either the model becoming larger (as the size is directly related to u^{max}) or that the relative differences from solution to solution is smaller when u^{max} is small. The relative performance differences, however, are marginal, showing that even though model (EXP) is pseudo-polynomial in u^{max}, **BCP** is able to scale for large preference ranges.

6.4. Boundary effects on α

When considering the single-objective variants of the TFP, when only total intra-group utility or stability, respectively, are of interest, a slight emphasis on the other objective leads to significant improvements. For $\alpha = 1$, maximizing utility only, a perturbation to $\alpha = 0.99$ results in slightly longer solution times but provides solutions with the same total utility but slightly better stability. When $\alpha = 0$, and one is concerned only with stability, perturbing to $\alpha = 0.01$ results in some solutions with even better stability within time limits, due to computational efficiency gains.

Focusing first on the $\alpha = 1$ case, we restrict attention to the 197 instances (out of 220) solved to optimality by **BCP** with $\alpha = 1$ and $\alpha = 0.99$. For every such instance, the total utility at the optimal solution is the same, but the maximum uplift can be significantly different. In 26 of the instances (13.2%) a better solution with respect to maximum uplift is identified, with reduction in maximum uplift ranging from 3.84% to 100%, and an average of 35.2%. For three instances in particular, the reported optimal solution left a maximum uplift of 8, 11, and 13, respectively under $\alpha = 1$, but with $\alpha = 0.99$, the reported optimal solution has the same total utility, but found an entirely stable solution (i.e., maximum uplift = 0, a reduction of 100%). This does come at a slight computational cost; the average solution time on these instances is 118 ($\alpha = 1$) and 147 ($\alpha = 0.99$) seconds, respectively. This added computational effort, however, is rewarded with solutions of strictly better quality. This suggests, for example, that recent research using BCPP (i.e., $\alpha = 1$) in sports management (e.g., Recalde et al. 2016) may find Pareto improved solutions using $\alpha = 0.99$.

The difference in algorithmic performance is even more apparent when comparing the solution quality and solution time under $\alpha = 0$ versus $\alpha =0.01$. First, only 73 instances are solved to optimality with $\alpha = 0$ versus 91 with $\alpha = 0.01$. More critically, despite having a theoretically worse optimal solution with respect to maximum uplift, in 131 of the 220 instances, **BCP** identifies a solution with smaller uplift when $\alpha = 0.01$ versus setting $\alpha = 0$, where the opposite occurs in only 3 instances. Furthermore, this reduction is often substantial. The average percent reduction in maximum uplift in the 131 instances where **BCP** found a solution with smaller maximum uplift is 74.5%. In 59 of the 131 instances, the reduction is 100%, meaning that a completely stable solution is found, whereas there are groups with positive uplift in the solution found with $\alpha = 0$ at 1800 seconds.

The reason for this surprising relative gain in solution quality can be attributed to the fact that solutions with high total intra-group utility will have relatively low instability, compared with randomly chosen groups. Without any consideration of total team utility in the objective function, it is challenging for the solver to distinguish between partitions of agents.

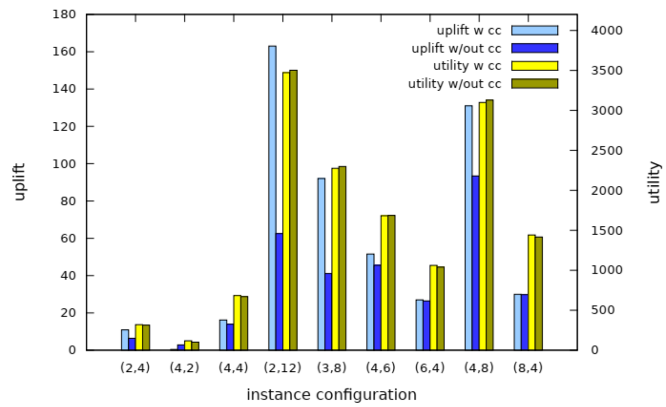


Figure 7: Average maximum uplift and average total market utility, with and without characteristic constraints.

Having a slight emphasis on total intra-group utility focuses the search for solutions, which turns out to be particularly effective under a column generation approach. **BCP**'s performance is only slightly hindered by the inclusion of characteristic constraints; BCP is able to solve 668 and 637 instances with and without characteristic constraints, respectively. Of the 617 instances that mutually solved with and without the constraints, the average solution time is 141.3 and 142.2 seconds, respectively, an insignificant difference.

6.5. The effect of characteristic constraints

Figure 7 provides a visual summary depicting the average of the maximum uplift and total utility, respectively, in the best solutions obtained by **BCP** with and without characteristic constraints. Clearly uplift is much more sensitive to adding characteristic constraints than utility. Indeed, the change in average utility is barely noticeable when constraints are added in any case, compared with the maximum uplift, which rises drastically in several market sizes, even more than doubling for some configurations. (Table 6 in the Appendix provides further detail.) Market designers should take heed; the addition of characteristic constraints may seem innocuous if only total utility is considered, but actually it has the potential to introduce a great deal of instability.

7. Comparison of BCP to DRAFT and OPOP

Here we compare **BCP** (with different α) to **DRAFT** and **OPOP** on three performance metrics: efficiency (measured by average individual utility), inequity (measured by the range of individual and team utility), and instability (measured by individual and coalitional uplift). **DRAFT** and **OPOP** (see §2.4) were not designed to handle characteristic constraints and may result in infeasibility if applied naively. (A team may for example over-consume a constrained characteristic, resulting in infeasibility for another team, even if they conscientiously pick to maintain their own feasibility throughout. Stated differently, given the NP-hardness of the problem, locally greedy algorithms will fail to maintain global feasibility.) The results of this section are therefore limited to an investigation of the $|Q| = 0$ subset of instances.

7.1. Efficiency

A clustered bar chart provides efficiency comparisons of **BCP**, **DRAFT**, and **OPOP** in Figure 8. For each agent in each outcome, we divide agent i's resulting utility by i's maximum realizable utility U(i). For each algorithm, Figure 8 reports the average of this percentage utility over all solutions for each generation scheme.

Our results show that **BCP** solutions have superior average utility across all sets of instances. This superiority is more obvious for data generation schemes with independent (**G3**) and affinity preferences (**G4**) and larger α. (The latter is to be expected as the weight of efficiency in the objective increases.) The significant common-value components of **G1** and **G2** result in a larger degree of "necessary disappointment" in which it is impossible for every agent to achieve her first best possible team. **BCP** tends to get closer to maximum satisfaction with the outcome of the mechanism. Perhaps surprisingly, even when the focus on efficiency is low (i.e., when $\alpha = 0.01$) **BCP** still provides more efficient solutions than the heuristic formats **DRAFT** and **OPOP**.

7.2. Inequity

Intuitively, having some agents with very high utility while other agents have very low utility can be problematic and viewed as inequitable or unfair. This can be measured and compared on the individual level and on the team level. For percentage-based comparisons, we define the individual inequity as the best individual utility minus the worst individual utility in a solution, divided by the average of U(*i*). Similarly, we define the team inequity as the best team utility minus the worst team utility in a solution divided by the average ideal-team utility, defined for each i as the maximum utility of any coalition including i. In Figure 9 we report the average individual inequity and team inequity over all instances in each generation scheme for each of the five algorithms. A taller bar indicates a worse solution on the measure of inequity.

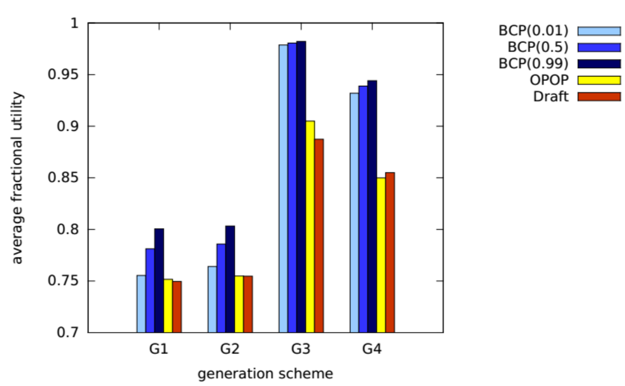


Figure 8: Average solution utility as a fraction of maximum realizable utility.

Clearly, the preference generation format can greatly impact the ranking of mechanisms on this measure of allocative inequity. The left chart in Figure 9 indicates that for data generation schemes G3 and G4, the individual inequity is typically smaller for **BCP** (regardless of α), and thus its results may be perceived as fairer. However, for data generation schemes with stronger common-value effects (G1 and G2), only **BCP** with α = 0.99 results in more equitable solutions than those found by **OPOP** and **DRAFT**.

Indeed, this relationship is among our more interesting findings; the focus on stability (inherent as α decreases) results in less equitable solutions when there is wide agreement on the "top of the market" as in **G1** and **G2**. Stated differently, in order to reduce the tendency of groups to want to break away and form their own teams (when the focus is stability), one will have to generate some teams with more "top" agents together, resulting in a wider satisfaction gap (inequity) in the market. This can only occur when the underlying preferences reflect a certain degree of agreement on which agents comprise the top of the market (e.g., in **G1** and **G2**). When agents tend to prefer to gather together based on their own affinity for those of similar characteristics (**G4**), this effect is reduced, and it nearly disappears when preferences are independent (**G3**).

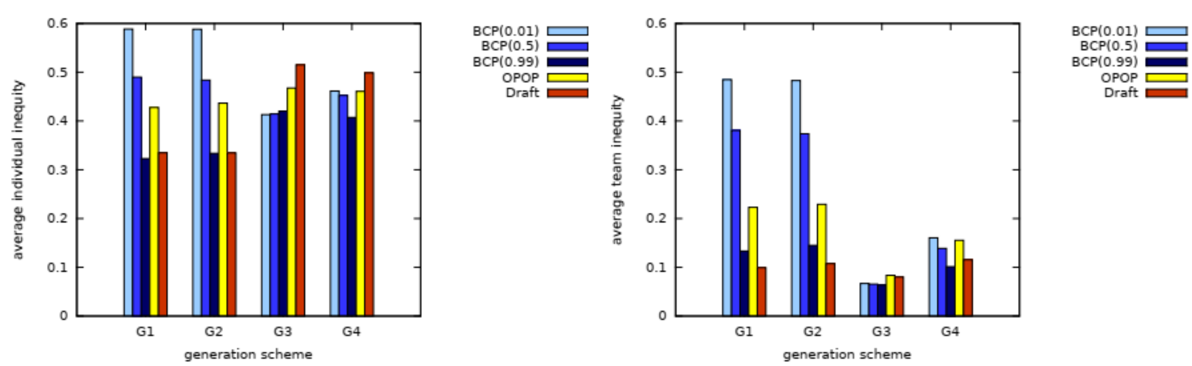


Figure 9: Average inequity (percentage utility range), averaged over individuals (left) and teams (right).

This effect is even more pronounced in the right half of Figure 9. Our results show that for data generation schemes with independent and affinity preferences (G3 and G4), **BCP** with α = 0.99 achieves more equitable teams than other algorithms. However, for data generation schemes **G1** and **G2**, **DRAFT** achieves the most equitable solutions across teams. Note that BCP is not directly minimizing inequity, and in particular when we focus only on minimizing the maximum uplift (α = 0.01), the team inequity is more obvious. **DRAFT** tends toward teams of equal satisfaction, spreading out "top picks" as would be expected; **BCP** with α = 0.01 tends to allow "top picks" to clump together in order to reduce their desire to leave the market.

7.3. Instability

To compare team formation outcomes across instances we again need a metric scaled to 1. Let individual instability be defined as the maximum uplift an agent can gain by defecting to a hypothetical uplift team, divided by her utility within her current team. Team instability is the maximum uplift of a team containing each i divided by the utility of her current team. These provide a relative measure of how much an individual or a team is incentivized to defect given the current team assignment. Figure 10 provides two plots depicting the average individual (left) and average coalition uplift (right), averaged over all *i*.

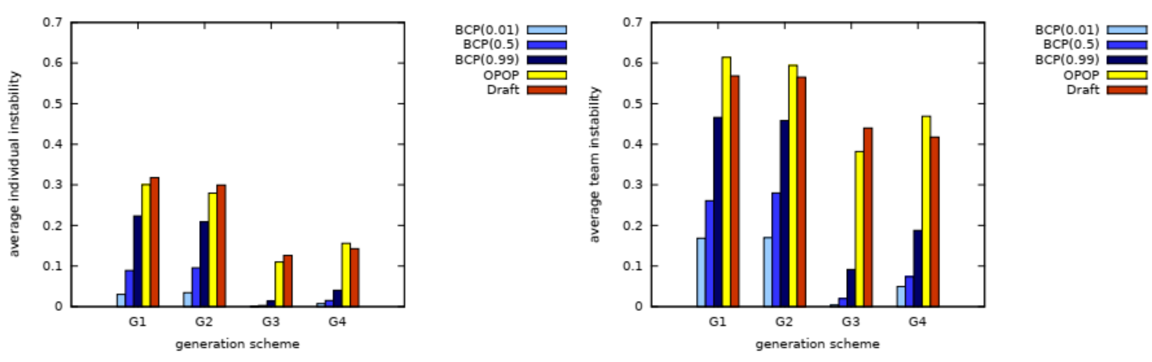


Figure 10: Average individual instability (left) and average team instability (right)

These charts show that regardless of α, the solutions obtained by **BCP** are more stable, having lower average and maximum individual uplifts. Moreover, the average individual benefit of forming a coalition in the solutions obtained by **BCP** with $\alpha = 0.01$ is less than 4% of their current benefit, while in **OPOP** and **DRAFT**, it can be on average as high as 30% of their current benefit. Also, note that the maximum uplift is significantly lower in data generation schemes **G3** and **G4** in comparison to data **G1** and **G2**. (Though not shown, the relative rankings do not change if we focus on maximum uplift as opposed to average.)

8. Conclusion and Future Work

In this paper we investigate computational models for the team formation problem. In order to balance intra-team utility with solution stability, we formulated a bi-level binary optimization model and developed a branch- cut-and-price solution algorithm. The pseudo-polynomial approach to the bi-level problem is itself an interesting contribution, and we detailed how to implement the resulting algorithm, which still has to manage an exponential number of variables and constraints, demanding advanced computational methods, which we outline. Experimental results indicated that the proposed algorithm BCP is particularly effective at finding high-quality solutions quickly. Stability as an objective to be optimized over remains a particularly challenging computational problem, but we have shown a promising new approach.

Our results also indicate that ignoring stability can result in inferior solutions when Pareto improvements are available. Because a defection-based measure of instability in this context points to opportunities for profitable group deviation (i.e., strategic manipulation), these results will have practical implications. We showed that heuristic algorithms from the literature may be leaving efficiency and stability gains on the table, as might have been expected. Yet heuristics based on draft principles may be common in practice; we would argue that this has much to do with the inherent computational difficulty of optimization approaches, which we have shown can be mitigated by algorithms like BCP. We have shown that our new methods can improve total satisfaction, and improve incentives via reduced instability measures.

While BCP with $\alpha = 0.99$ performed very well on equity as a measure of fairness, we found an interesting trade-off, in which the market maker must in some cases accept inter-agent inequity in pursuit of stability (as α goes to zero), in particular it would seem, where common value is the primary driver of preferences. This discrepancy of inequity performance across preference models proved interesting, and we expect future studies to shed more light on how different models or descriptions of preferences affect the performance of various algorithms for team formation.

References

David J. Abraham, P´eter Bir´o, and David F. Manlove. *\Almost Stable" Matchings in the Roommates Problem*, pages 1{14. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-32208-5. doi: 10.1007/11671411*n* 1. URL https://doi.org/10.1007/11671411_1.

Rakesh Agrawal, Behzad Golshan, and Evimaria Terzi. Grouping students in educational settings. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Minining (KDD)*. ACM, August 2014. URL https://www.microsoft.com/en-us/research/publication/grouping-students-in-educational-settings/.

Haris Aziz, Felix Brandt, and Hans Georg Seedig. Stable partitions in additively separable hedonic games. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 1* , AAMAS '11, pages 183{190, Richland, SC, 2011. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 0-9826571-5-3, 978-0-9826571-5-7. URL http://dl.acm.org/citation.cfm?id=2030470.2030497.

Helen Baron. Strengths and limitations of ipsative measurement. *Journal of Occupational and Organizational Psychology*, 69(1):49{56, 1996.

J. Bhasker and Tariq Samad. The clique-partitioning problem. *Computers & Mathematics with Applications*, 22(6):1 { 11, 1991. ISSN 0898-1221. doi: https://doi.org/10.1016/0898-1221(91)90001-K. URL http://www.sciencedirect.com/science/article/pii/089812219190001K.

P´eter Bir´o, Elena I narra, and Elena Molis. A new solution concept for the roommate problem: Q-stable matchings. *Mathematical Social Sciences*, 79 (Supplement C):74 {82, 2016. ISSN 0165-4896. doi: https://doi.org/10.1016/j.mathsocsci.2015.12.001. URL http://www.sciencedirect.com/science/article/pii/S0165489615001067.

Michael A. Campion, Gina J. Medsker, and A. Catherine Higgs. Relations between work group characteristics and effectiveness: implications for designing effective work groups. *Personnel Psychology*, 46(4):823{847, 1993. ISSN 1744- 6570. doi: 10.1111/j.1744-6570.1993.tb01571.x. URL http://dx.doi.org/10.1111/j.1744-6570.1993.tb01571.x.

Katar´ına Cechl´arov´a and Tam´as Fleiner. On a generalization of the stable roommates problem. *ACM Trans. Algorithms*, 1(1):143{156, July 2005. ISSN 1549-6325. doi: 10.1145/1077464.1077474. URL http://doi.acm.org/10.1145/1077464.1077474.

Rong-Chang Chen, Shih-Ying Chen, Jyun-You Fan, and Yen-Ting Chen. Grouping partners for cooperative learning using genetic algorithm and social network analysis. *Procedia Engineering*, 29(Supplement C):3888 { 3893, 2012. ISSN 1877-7058. doi: https://doi.org/10.1016/j.proeng.2012.01.589. URL http://www.sciencedirect.com/science/article/pii/S1877705812005991. 2012 International Workshop on Information and Electronics Engineering.

Aytek Erdil and Haluk Ergin. What's the matter with tie-breaking? Improving efficiency in school choice. *American Economic Review*, 98(3):669{89, 2008.

F. Farhadi, M. Sorkhi, S. Hashemi, and A. Hamzeh. An effective expert team formation in social networks based on skill grading. In *2011 IEEE 11th International Conference on Data Mining Workshops*, pages 366{372, Dec 2011. doi: 10.1109/ICDMW.2011.28.

David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9{15, 1962.

Richard G Gardner, T Brad Harris, Ning Li, Bradley L Kirkman, and John E Mathieu. Understanding it depends in organizational research: A theorybased taxonomy, review, and future research agenda concerning interactive and quadratic relationships. *Organizational Research Methods*, page 1094428117708856, 2017.

Dan Gusfield and Robert W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, Cambridge, MA, USA, 1989. ISBN 0-262-07118-5.

Jimmy H. Gutierrez, Csar A. Astudillo, Pablo Ballesteros-Perez, Daniel MoraMelia, and Alfredo Candia-Vejar. The multiple team formation problem using sociometry. *Computers & Operations Research*, 75(Supplement C):150 { 162, 2016. ISSN 0305-0548. doi: https://doi.org/10.1016/j.cor.2016.05.012. URL http://www.sciencedirect.com/science/article/pii/S0305054816301198.

Robert W Irving. An efficient algorithm for the stable roommates problem. *Journal of Algorithms*, 6(4):577 { 595, 1985. ISSN 0196-6774. doi: https://doi.org/10.1016/0196-6774(85)90033-1. URL http://www.sciencedirect.com/science/article/pii/0196677485900331.

Robert W. Irving. Stable marriage and indifference. *Discrete Applied Mathematics*, 48(3):261{272, 1994. doi: 10.1016/0166-218X(92)00179-P. URL https://doi.org/10.1016/0166-218X(92)00179-P.

Kazuo Iwama and Shuichi Miyazaki. A survey of the stable marriage problem and its variants. In *Proceedings of the International Conference on Informatics Education and Research for Knowledge-Circulating Society (Icks 2008)*, ICKS '08, pages 131{136, Washington, DC, USA, 2008. IEEE Computer Society. ISBN 978-0-7695-3128-1. doi: 10.1109/ICKS.2008.7. URL https://doi. org/10.1109/ICKS.2008.7.

Ricardo M. Lima and Ignacio E. Grossmann. On the solution of nonconvex cardinality boolean quadratic programming problems: a computational study. *Computational Optimization and Applications*, 66(1):1{37, Jan 2017. ISSN 1573-2894. doi: 10.1007/s10589-016-9856-7. URL https://doi.org/10. 1007/s10589-016-9856-7.

John E Mathieu, Scott I Tannenbaum, Michael R Kukenberger, Jamie S Donsbach, and George M Alliger. Team role experience and orientation: A measure and tests of construct validity. *Group & Organization Management*, 40(1):6-34, 2015.

Abraham Othman, Tuomas Sandholm, and Eric Budish. Finding approximate competitive equilibria: Efficient and fair course allocation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 873{880. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

Joana Pais and Agnes Pint´er. School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior*, 64(1):303-328, 2008.

Patrick Prosser. *Stable Roommates and Constraint Programming*, pages 15-28. Springer International Publishing, Cham, 2014. ISBN 978-3-319-07046-9. doi: 10.1007/978-3-319-07046-9*n* 2. URL https://doi.org/10.1007/978-3-319-07046-9_2.

Diego Recalde, Daniel Sever´ın, Ramiro Torres, and Polo Vaca. *Balanced Partition of a Graph for Football Team Realignment in Ecuador*, pages 357-368. Springer International Publishing, Cham, 2016. ISBN 978-3-319-45587- 7. doi: 10.1007/978-3-319-45587-7*n* 31. URL https://doi.org/10.1007/978-3-319-45587-7_31.

Alvin E Roth. The national residency matching program as a labor market. *JAMA*, 275(13):1054{1056, 1996. Alvin E Roth, Tayfun S¨onmez, and M Utku Unver. Kidney exchange. ¨ *The Quarterly Journal of Economics*, 119(2):457-488, 2004.

Shao-Chin Sung and Dinko Dimitrov. Computational complexity in additive hedonic games. *European Journal of Operational Research*, 203(3):635{639, 2010. Jimmy J. M. Tan. A maximum stable matching for the roommates problem. *BIT Numerical Mathematics*, 30(4):631-640, Dec 1990. ISSN 1572-9125. doi: 10. 1007/BF01933211. URL https://doi.org/10.1007/BF01933211.

The Myers & Briggs Foundation. MBTI, 2018. URL http://www.myersbriggs. org/.

Martin Van der Linden, Greg Leo, Jian Lou, Yevgeniy Vorobeychik, and Myrna H Wooders. Matching soulmates. 2016. Ouri Wolfson and Jane Lin. Fairness versus optimality in ridesharing. In *Mobile Data Management (MDM), 2017 18th IEEE International Conference on*, pages 118-123. IEEE, 2017.

Mason Wright and Yevgeniy Vorobeychik. Mechanism design for team formation. *CoRR*, abs/1501.00715, 2015. URL http://arxiv.org/abs/1501.00715

Appendix

Table 2 through Table 5 provide detailed solution statistics on **G1**, **G2**, **G3**, and **G4**, respectively. The first four columns report the number of agents, the number of groups, the maximum pairwise utility, and α. The next three sets of four columns report solution statistics for **EXP**, **BC**, and **BCP**, respectively. In particular, for each algorithm and each configuration, we report the number of instances solved in 1800 seconds (n^s), the average solution time over those n^2 instances (time), the number of instances that have a provable optimality gap n^g (i.e., for **EXP** the number of instances that don't hit memory limits and for **BCP** the number of instances for which the root node was solved and an initial upper bound is proven), and finally the average gap over those n^g instances. The last two columns report the average total utility (u^{best}) and maximum uplift (r^{best}) for the best solution identified by **BCP** for that configuration.

Table 6 provides details on the effect of including characteristic constraints (considering only those instances with u^{max} = 25). For each instance configuration and for $\alpha \in \{0.01, 0.5, .99\}$, the table reports, for the best solution identified by **BCP**, the average uplift ($\bar{r}$) and average total utility ($\bar{u}$) with and without characteristic constraints, for the instances from each generation scheme, in sequence.

Table 2: Detailed solution statistics for **G1**

				EXP				BC				BCP					
n	m	$u^{\max}$	α	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	u^{best}	r^{best}
8	2	25	0	5	0.07	5	0	5	0.01	5	0	5	0.09	5	0	302	0
			0.01	5	0.45	5	0	5	0.92	5	0	5	2.61	5	0	304.2	0
			0.5	5	0.49	5	0	5	0.44	5	0	5	2.74	5	0	304.2	0
			0.99	5	0.09	5	0	5	0.92	5	0	5	0.21	5	0	319	36.6
			1	5	0.04	5	0	5	0.43	5	0	5	0.19	5	0	319	36.6
	4	25	0	5	0.01	5	0	5	0.74	5	0	5	0.02	5	0	109	0
			0.01	5	0.01	5	0	5	0.73	5	0	5	0.13	5	0	109	0
			0.5	5	0.01	5	0	5	1.1	5	0	5	0.12	5	0	110.6	1
			0.99	5	0.01	5	0	5	1.37	5	0	5	0.03	5	0	113.4	9.8
			1	5	0.01	5	0	5	1.24	5	0	5	0.03	5	0	113.4	9.8
16	4	25	0	5	481.8	5	0	1	1607	5	23.2	0	-	5	20.4	612.8	20.4
			0.01	5	391.9	5	0	0	-	5	16.48	1	832	5	3.51	633.4	3.4
			0.5	5	528.2	5	0	0	-	5	216.1	1	831.7	5	4.39	651.8	12.8
			0.99	5	4.05	5	0	0	-	5	371	5	4.42	5	0	666.6	47
			1	5	0.62	5	0	0	-	5	346.8	5	3.83	5	0	666.6	47.6
24	2	25	0	0	-	0	-	5	0.22	5	0	5	1.14	5	0	3313	0
			0.01	0	-	0	-	0	-	5	4.2	0	-	4	1.01	3313	0
			0.5	0	-	0	-	0	-	5	225.3	0	-	3	47.61	3313	0
			0.99	0	-	0	-	0	-	5	304.7	3	1346	3	0	3429	370.4
			1	0	-	0	-	0	-	5	367.2	0	-	0	-	3429	370.4
	3	25	0	0	-	0	-	5	0.29	5	0	5	0.69	5	0	2094	0
			0.01	0	-	0	-	0	-	5	17.63	0	-	5	1.48	2094	0
			0.5	0	-	0	-	0	-	5	864.1	0	-	5	73.76	2094	0
			0.99	0	-	0	-	0	-	5	1605	5	1124	5	0	2246	209.2
			1	0	-	0	-	0	-	5	1650	5	633.7	5	0	2246	209.2
	4	25	0	0	-	0	-	0	-	5	124.6	0	-	5	49.6	1487	49.6
			0.01	0	-	0	-	0	-	5	114.1	0	-	5	45.56	1510	44.8
			0.5	0	-	0	-	0	-	5	709.7	0	-	5	40.6	1625	75.8
			0.99	0	-	0	-	0	-	5	1283	5	228.8	5	0	1637	126.6
			1	5	130	5	0	0	-	5	1289	5	184.1	5	0	1637	126.6
	6	5	0	0	-	5	4.8	0	-	5	8.2	0	-	5	8.8	179.4	8.8
			0.01	0	-	5	6.74	0	-	5	11.06	0	-	5	4.43	200.6	4.4
			0.5	1	1122	5	1.2	0	-	5	86.3	0	-	5	2.77	209.6	7.2
			0.99	5	27.35	5	0	0	-	5	160.7	5	50.38	5	0	210.6	9
			1	5	1.67	5	0	0	-	5	160.2	5	17.98	5	0	210.6	10.4
		25	0	0	-	5	42.2	0	-	5	63.8	0	-	5	33.2	903.6	33.2
			0.01	0	-	5	31.32	0	-	5	49.34	0	-	5	20.35	999	20.4
			0.5	0	-	5	20.9	0	-	5	363.9	0	-	5	16.3	1010	28.4
			0.99	5	50.38	5	0	0	-	5	676.5	5	35.75	5	0	1022	56.2
			1	5	4.99	5	0	0	-	5	681.8	5	31.1	5	0	1022	56.2
		100	0	0	-	5	250.4	0	-	5	245	0	-	5	133.8	3620	133.8
			0.01	0	-	5	131.1	0	-	5	217.9	0	-	5	69.01	3862	67.4
			0.5	0	-	5	119.1	0	-	5	1565	0	-	5	74.54	4063	129.2
			0.99	5	508.7	5	0	0	-	5	2881	5	107.4	5	0	4096	229.6
			1	5	14.98	5	0	0	-	5	2899	5	82.16	5	0	4096	229.6
32	4	25	0	0	-	0	-	1	0.85	5	207.2	1	1.29	5	65	2777	65
			0.01	0	-	0	-	0	-	5	172.1	0	-	0	-	2769	66.4
			0.5	0	-	0	-	0	-	5	1325	0	-	0	-	2984	171.4
			0.99	0	-	0	-	0	-	5	2413	0	-	0	-	3028	240.8
			1	0	-	0	-	0	-	5	2442	0	-	0	-	3028	233.4
	8	25	0	0	-	5	63.6	0	-	5	65	0	-	5	41.6	1199	41.6
			0.01	0	-	5	39.59	0	-	5	69.91	0	-	5	27.58	1298	27
			0.5	0	-	5	30.3	0	-	5	550.2	0	-	5	25.56	1378	45.8
			0.99	5	465.7	5	0	0	-	5	1024	5	164.7	5	0	1385	64.4
			1	5	15.7	5	0	0	-	5	1036	5	109.5	5	0	1385	65.2

Table 3: Detailed solution statistics for **G2**

				EXP				BC				BCP					
n	m	$u^{\max}$	α	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	u^{best}	r^{best}
8	2	25	0	5	0.07	5	0	5	0.12	5	0	5	0.49	5	0	300.8	0
			0.01	5	0.23	5	0	5	0.52	5	0	5	3.41	5	0	305	0
			0.5	5	0.36	5	0	5	0.42	5	0	5	3.05	5	0	305	0
			0.99	5	0.09	5	0	5	0.62	5	0	5	0.17	5	0	315.8	29.6
			1	5	0.04	5	0	5	0.45	5	0	5	0.16	5	0	315.8	29.6
	4	25	0	5	0.01	5	0	5	0.72	5	0	5	1.02	5	0	106.6	2
			0.01	5	0.01	5	0	5	0.72	5	0	5	0.97	5	0	106.6	2
			0.5	5	0.01	5	0	5	0.89	5	0	5	0.3	5	0	110.4	2.4
			0.99	5	0.01	5	0	5	0.96	5	0	5	0.04	5	0	113.2	8.6
			1	5	0.01	5	0	5	0.95	5	0	5	0.03	5	0	113.2	9.2
16	4	25	0	5	698.3	5	0	1	0.08	5	26.8	1	0.22	5	16.4	612.4	16.4
			0.01	4	433.5	5	3.8	0	-	5	17.43	1	449.1	5	6.67	636.2	6.6
			0.5	5	520.3	5	0	0	-	5	211.9	1	528.7	5	5.13	652.8	14
			0.99	5	3.1	5	0	0	-	5	376.4	5	5.66	5	0	668	44.4
			1	5	0.62	5	0	0	-	5	364.6	5	4.47	5	0	668	44.4
24	2	25	0	0	-	0	-	5	0.22	5	0	5	1.18	5	0	3305	0
			0.01	0	-	0	-	0	-	5	4.43	0	-	2	0.95	3305	0
			0.5	0	-	0	-	0	-	5	208.9	0	-	2	46.25	3305	0
			0.99	0	-	0	-	0	-	5	367.3	1	1315	2	1.07	3419	369.6
			1	0	-	0	-	0	-	5	355	0	-	0	-	3419	369.6
	3	25	0	0	-	0	-	4	0.27	5	48.4	4	0.68	5	9.8	2101	9.8
			0.01	0	-	0	-	0	-	5	29.77	0	-	5	19.67	2096	18.4
			0.5	0	-	0	-	0	-	5	856.2	0	-	5	71.42	2124	27.6
			0.99	0	-	0	-	0	-	5	1633	5	834.5	5	0	2247	202.6
			1	0	-	0	-	0	-	5	1644	5	827.7	5	0	2247	202.6
	4	25	0	0	-	0	-	0	-	5	128.2	0	-	5	34.8	1502	34.8
			0.01	0	-	0	-	0	-	5	105.2	0	-	4	35.59	1525	43.6
			0.5	0	-	0	-	0	-	5	668.3	0	-	5	44.38	1622	74.8
			0.99	0	-	0	-	0	-	5	1200	5	301.8	5	0	1642	118
			1	5	138.6	5	0	0	-	5	1223	5	172.3	5	0	1642	118
	6	5	0	0	-	5	4	0	-	5	7.8	0	-	5	7.2	183.2	7.2
			0.01	0	-	5	5.5	0	-	5	11.1	0	-	5	5.41	201	5.4
			0.5	2	1462	5	1.8	0	-	5	71.1	0	-	5	2.85	208.6	6.8
			0.99	5	27.31	5	0	0	-	5	129.4	5	39.32	5	0	210.2	9.8
			1	5	1.48	5	0	0	-	5	130.8	5	15.19	5	0	210.2	11.8
		25	0	0	-	5	37.6	0	-	5	61.8	0	-	5	39	904	39
			0.01	0	-	5	32.72	0	-	5	54.94	0	-	5	22.39	969.4	22.2
			0.5	0	-	5	21.9	0	-	5	366.2	0	-	5	18.33	1010	37.4
			0.99	5	56.24	5	0	0	-	5	673.6	5	45.67	5	0	1017	58.4
			1	5	4.67	5	0	0	-	5	684.2	5	27.76	5	0	1017	63.6
		100	0	0	-	5	229.8	0	-	5	252	0	-	5	133.6	3673	133.6
			0.01	0	-	5	134.6	0	-	5	230.1	0	-	5	83.74	3964	83.4
			0.5	0	-	5	104.3	0	-	5	1616	0	-	5	76.22	4074	144
			0.99	5	150.7	5	0	0	-	5	2960	5	90.41	5	0	4098	218.2
			1	5	14.22	5	0	0	-	5	2929	5	50.88	5	0	4098	218.2
32	4	25	0	0	-	0	-	3	0.85	5	92.8	3	1.39	5	21.6	2792	21.6
			0.01	0	-	0	-	0	-	5	107.2	0	-	0	-	2792	21.6
			0.5	0	-	0	-	0	-	5	1294	0	-	0	-	2963	172
			0.99	0	-	0	-	0	-	5	2418	0	-	0	-	3011	247.4
			1	0	-	0	-	0	-	5	2434	0	-	0	-	3019	236.4
	8	25	0	0	-	5	65.8	0	-	5	59.2	0	-	5	39.4	1211	39.4
			0.01	0	-	5	40.79	0	-	5	75.04	0	-	5	22.3	1317	21.8
			0.5	0	-	5	31.8	0	-	5	527.6	0	-	5	22	1376	34.8
			0.99	5	594.2	5	0	0	-	5	954.2	5	129.2	5	0	1389	64
			1	5	15.15	5	0	0	-	5	971.8	5	105.2	5	0	1389	64

Table 4: Detailed solution statistics for **G3**

				EXP				BC				BCP					
n	m	$u^{\max}$	α	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	u^{best}	r^{best}
8	2	25	0	5	0.29	5	0	5	1.26	5	0	5	0.35	5	0	325.2	0
			0.01	5	0.09	5	0	5	1.16	5	0	5	0.16	5	0	329.8	0
			0.5	5	0.09	5	0	5	0.75	5	0	5	0.16	5	0	329.8	0
			0.99	5	0.09	5	0	5	0.98	5	0	5	0.16	5	0	330	2.8
			1	5	0.04	5	0	5	0.94	5	0	5	0.17	5	0	330	2.8
	4	25	0	5	0.01	5	0	5	0.97	5	0	5	0.13	5	0	126.6	1
			0.01	5	0.01	5	0	5	0.96	5	0	5	0.2	5	0	126.6	1
			0.5	5	0.01	5	0	5	1.1	5	0	5	0.08	5	0	127.8	1.6
			0.99	5	0.01	5	0	5	1.31	5	0	5	0.03	5	0	128.8	4
			1	5	0.01	5	0	5	1.47	5	0	5	0.03	5	0	128.8	6.6
16	4	25	0	2	832.6	5	7.6	0	-	5	27.8	3	327.8	5	7	679.4	7
			0.01	5	134	5	0	0	-	5	5.1	5	36.74	5	0	713.4	0
			0.5	5	39.15	5	0	0	-	5	213	5	28.86	5	0	718	4.6
			0.99	5	3.29	5	0	0	-	5	439.2	5	2.32	5	0	720.8	11.6
			1	5	0.71	5	0	0	-	5	413.8	5	2.03	5	0	720.8	11.6
24	2	25	0	0	-	0	-	4	92.82	5	19.8	5	309	5	0	3382	0
			0.01	0	-	0	-	0	-	5	4.11	4	585.2	4	0	3516	0
			0.5	0	-	0	-	1	1561	5	135	4	418	4	0	3516	0
			0.99	0	-	0	-	1	1706	5	218.6	4	428.6	4	0	3516	0
			1	0	-	0	-	2	1459	5	215.2	4	432.6	4	0	3516	0
	3	25	0	0	-	0	-	0	-	5	113.4	0	-	5	86.6	2186	86.6
			0.01	0	-	0	-	0	-	5	63.1	5	112.8	5	0	2347	0
			0.5	0	-	0	-	0	-	5	946.5	5	114.7	5	0	2347	0
			0.99	0	-	0	-	0	-	5	1829	5	115.7	5	0	2347	0
			1	0	-	0	-	0	-	5	1878	5	114.2	5	0	2347	0
	4	25	0	0	-	0	-	0	-	5	85	0	-	5	72.2	1584	72.2
			0.01	0	-	0	-	0	-	5	65.9	5	61.93	5	0	1756	0
			0.5	0	-	0	-	0	-	5	698.3	5	56.21	5	0	1756	0
			0.99	0	-	0	-	0	-	5	1322	5	39.82	5	0	1757	6.6
			1	5	131.2	5	0	0	-	5	1336	5	45.35	5	0	1757	6.6
	6	5	0	0	-	5	7.4	0	-	5	7.2	0	-	5	8.4	197	8.4
			0.01	0	-	5	2.37	0	-	5	8.66	0	-	5	1.19	224.6	1.2
			0.5	5	215.1	5	0	0	-	5	79.4	5	601.9	5	0	229	3.2
			0.99	5	12.81	5	0	0	-	5	148.8	5	13.55	5	0	229.8	5.4
			1	5	1.46	5	0	0	-	5	151.4	5	8.43	5	0	229.8	6
		25	0	0	-	5	40	0	-	5	52.4	0	-	5	42.6	955	42.6
			0.01	2	833.7	5	5.76	0	-	5	44.14	5	543.8	5	0	1118	0
			0.5	4	746.6	5	0.4	0	-	5	415	5	184.2	5	0	1125	4.2
			0.99	5	58.2	5	0	0	-	5	756.8	5	27.27	5	0	1128	25.6
			1	5	4.87	5	0	0	-	5	772.6	5	15.77	5	0	1128	25.8
		100	0	0	-	5	239.4	0	-	5	217.6	0	-	5	178.2	3840	178.2
			0.01	2	88.69	5	38.18	0	-	5	197.2	5	241.4	5	0	4489	0
			0.5	2	75.23	3	4.33	0	-	5	1724	5	195.1	5	0	4489	0
			0.99	5	121.4	5	0	0	-	5	3195	5	41.02	5	0	4503	54.6
			1	5	15.76	5	0	0	-	5	3187	5	34.15	5	0	4503	58.6
32	4	25	0	0	-	0	-	0	-	5	127.6	0	-	5	127.8	2869	127.8
			0.01	0	-	0	-	0	-	5	128.6	3	921.6	5	0.02	3204	0
			0.5	0	-	0	-	0	-	5	1371	4	976.8	5	0.33	3205	0
			0.99	0	-	0	-	0	-	5	2610	5	980.2	5	0	3208	16.4
			1	0	-	0	-	0	-	5	2623	4	669.6	5	0.31	3208	23.4
	8	25	0	0	-	5	56.8	0	-	5	63.6	0	-	5	60	1228	60
			0.01	0	-	5	18.65	0	-	5	67.79	0	-	5	1.19	1530	1
			0.5	1	1244	5	7.5	0	-	5	568.8	2	868.6	5	1.57	1550	12.6
			0.99	5	210	5	0	0	-	5	1089	5	68.13	5	0	1554	20.6
			1	5	15.06	5	0	0	-	5	1084	5	55.95	5	0	1554	27

Table 5: Detailed solution statistics for **G4**

				EXP				BC				BCP					
n	m	$u^{\max}$	α	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	n^s	$\overline{\text{time}}$	n^g	$\overline{\text{gap}}$	u^{best}	r^{best}
8	2	25	0	5	0.19	5	0	5	1.29	5	0	5	0.36	5	0	324.6	0
			0.01	5	0.09	5	0	5	0.78	5	0	5	0.37	5	0	332.2	0
			0.5	5	0.11	5	0	5	0.39	5	0	5	0.35	5	0	332.2	0
			0.99	5	0.08	5	0	5	0.46	5	0	5	0.16	5	0	335.8	7.4
			1	5	0.04	5	0	5	0.59	5	0	5	0.19	5	0	335.8	7.4
	4	25	0	5	0.01	5	0	5	0.51	5	0	5	0.3	5	0	121.6	0.6
			0.01	5	0.01	5	0	5	0.47	5	0	5	0.28	5	0	121.6	0.6
			0.5	5	0.01	5	0	5	0.48	5	0	5	0.06	5	0	124.8	1.6
			0.99	5	0.01	5	0	5	0.41	5	0	5	0.03	5	0	124.8	1.6
			1	5	0.01	5	0	5	0.38	5	0	5	0.03	5	0	124.8	5.4
16	4	25	0	5	600.2	5	0	0	-	5	35.4	0	-	5	32.4	653	32.4
			0.01	5	219.9	5	0	0	-	5	12.19	5	413.6	5	0	704	0
			0.5	5	38.45	5	0	0	-	5	154.7	5	71.4	5	0	710.4	1.6
			0.99	5	4.35	5	0	0	-	5	283.3	5	3.35	5	0	723.2	21.8
			1	5	0.64	5	0	0	-	5	276.2	5	3.18	5	0	723.2	21.8
24	2	25	0	0	-	0	-	0	-	5	224	1	383.4	5	148.2	3397	148.2
			0.01	0	-	0	-	2	1344	5	5.56	4	388.8	4	0	3678	0
			0.5	0	-	0	-	4	1068	5	28.4	5	372.5	5	0	3678	0
			0.99	0	-	0	-	5	1024	5	0	4	324.8	4	0	3680	10.4
			1	0	-	0	-	5	973.1	5	0	5	329.2	5	0	3680	10.4
	3	25	0	0	-	0	-	0	-	5	185.2	0	-	5	143.2	2159	143.2
			0.01	0	-	0	-	0	-	5	108.3	4	225.2	5	0	2452	0
			0.5	0	-	0	-	0	-	5	806.1	4	219.9	5	0.12	2452	0
			0.99	0	-	0	-	0	-	5	1429	5	255.6	5	0	2462	35.6
			1	0	-	0	-	0	-	5	1536	5	79.09	5	0	2462	35.6
	4	25	0	0	-	0	-	0	-	5	115.4	0	-	5	100	1557	100
			0.01	0	-	0	-	0	-	5	93.32	3	240.8	5	5.63	1777	5.6
			0.5	0	-	0	-	0	-	5	567.3	3	201.7	5	3.78	1785	6.8
			0.99	0	-	0	-	0	-	5	1019	5	64.81	5	0	1797	43.8
			1	5	141.9	5	0	0	-	5	1021	5	73.92	5	0	1797	51.8
	6	5	0	0	-	5	5.6	0	-	5	8.2	0	-	5	9.4	192.8	9.4
			0.01	1	1657	5	3.15	0	-	5	9.26	0	-	5	2.77	221.4	2.8
			0.5	5	168.3	5	0	0	-	5	58.2	3	712.9	5	0.13	225.6	4.4
			0.99	5	22.44	5	0	0	-	5	111.4	5	25.75	5	0	225.6	4.4
			1	5	1.55	5	0	0	-	5	105.8	5	16.26	5	0	225.6	5.6
		25	0	0	-	5	53.8	0	-	5	61.4	0	-	5	52.2	959.8	52.2
			0.01	0	-	5	17.91	0	-	5	59.87	0	-	5	9.96	1087	9.8
			0.5	1	1769	5	6.5	0	-	5	345.1	0	-	5	4.35	1120	21.2
			0.99	5	39.38	5	0	0	-	5	618.7	5	39.06	5	0	1125	33.4
			1	5	4.94	5	0	0	-	5	635.6	5	30.87	5	0	1125	36
		100	0	0	-	5	258	0	-	5	254.6	0	-	5	204.4	3955	204.4
			0.01	0	-	5	89.05	0	-	5	237.3	0	-	5	6.59	4456	6
			0.5	0	-	4	35.75	0	-	5	1397	1	1232	5	11.38	4507	27.2
			0.99	5	124.3	5	0	0	-	5	2506	5	88.97	5	0	4540	105.2
			1	5	14.06	5	0	0	-	5	2563	5	106.9	5	0	4540	105.2
32	4	25	0	0	-	0	-	0	-	5	223.6	0	-	5	182.4	2908	182.4
			0.01	0	-	0	-	0	-	5	195.6	1	914.8	5	46.71	3329	47
			0.5	0	-	0	-	0	-	5	1238	1	1423	5	25.34	3338	42.6
			0.99	0	-	0	-	0	-	5	2317	5	437.2	5	0	3356	95.8
			1	0	-	0	-	0	-	5	2316	5	508.4	5	0	3356	95.8
	8	25	0	0	-	5	62.6	0	-	5	67.6	0	-	5	62.2	1256	62.2
			0.01	0	-	5	23.44	0	-	5	72.39	0	-	5	13.42	1482	13.2
			0.5	0	-	5	15.2	0	-	5	519.8	0	-	5	8.75	1515	20.4
			0.99	5	216.1	5	0	0	-	5	918	5	113.8	5	0	1519	32.4
			1	5	15.17	5	0	0	-	5	943.6	5	110.5	5	0	1519	33.8

Table 6: Comparing solutions with and without characteristic constraints

			G1				G2				G3				G4			
			W/out Cons		W/ Cons		W/out Cons		W/ Cons		W/out Cons		W/ Cons		W/out Cons		W/ Cons	
n	m	α	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$	$\bar{r}$	$\bar{u}$
8	2	0.01	0	304.2	4.6	303.6	0	305	11.4	304.2	0	329.8	0	324.2	0	332.2	3	315.2
		0.5	0	304.2	4.6	303.6	0	305	11.6	305	0	329.8	0	324.2	0	332.2	3	315.2
		0.99	36.6	319	41.4	316	29.6	315.8	37	312.4	2.8	330	5.2	325.8	7.4	335.8	8.6	316
	4	0.01	0	109	0	100.2	2	106.6	0	99.2	1	126.6	0	110.8	0.6	121.6	0	92.6
		0.5	1	110.6	0	100.2	2.4	110.4	0	99.2	1.6	127.8	0	110.8	1.6	124.8	0.2	93
		0.99	9.8	113.4	3.6	102.2	8.6	113.2	0	99.2	4	128.8	0.8	111	1.6	124.8	0.2	93
16	4	0.01	3.4	633.4	13.2	630.2	6.6	636.2	2.4	639.6	0	713.4	0	709	0	704	4.2	663.8
		0.5	12.8	651.8	19.2	642.8	14	652.8	19.4	656.8	4.6	718	1.8	711.4	1.6	710.4	8.6	674.6
		0.99	47	666.6	52.2	657.8	44.4	668	38.2	666.2	11.6	720.8	12.4	714.6	21.8	723.2	23.2	679.8
24	2	0.01	0	3312.8	303.8	3402.4	0	3305.2	314.6	3400.8	0	3516	0	3516	0	3677.6	0	3668.8
		0.5	0	3312.8	254.4	3402.6	0	3305.2	308	3405.6	0	3516	0	3516	0	3677.6	12.4	3672.8
		0.99	370.4	3428.8	370.4	3428.8	369.6	3418.6	369.6	3418.6	0	3516	0	3516	10.4	3680.4	22.8	3675.6
	3	0.01	0	2094.2	156.4	2207.4	18.4	2095.8	141.4	2214.6	0	2347.2	0	2347.2	0	2452.2	10.4	2385.2
		0.5	0	2094.2	150.8	2220.2	27.6	2124.4	150.4	2229.2	0	2347.2	0	2347.2	0	2452.2	10.6	2388
		0.99	209.2	2245.6	209.2	2245.6	202.6	2246.8	202.6	2246.8	0	2347.2	0	2347.2	35.6	2461.6	72.6	2404.6
	4	0.01	44.8	1509.6	81.6	1610.8	43.6	1525.4	73.4	1617.4	0	1755.8	0	1755.4	5.6	1777	2	1732
		0.5	75.8	1624.6	75.6	1621.2	74.8	1622.4	74.6	1624.6	0	1755.8	0	1755.4	6.8	1784.8	2	1732
		0.99	126.6	1636.8	130.6	1635.6	118	1641.6	118	1641.6	6.6	1757.2	10.8	1757.2	43.8	1797.4	50	1747.8
	6	0.01	20.4	999	27.2	995.2	22.2	969.4	24	985.8	0	1118.2	0	1102.6	9.8	1087.2	18.6	1034.6
		0.5	28.4	1009.8	29.2	1004	37.4	1010	33.8	1002.2	4.2	1124.8	4.2	1107.4	21.2	1120	24	1046.4
		0.99	56.2	1021.8	58.8	1016.2	58.4	1017.4	57.4	1011	25.6	1128.2	18.8	1113.8	33.4	1125.2	27.8	1047.4
32	4	0.01	66.4	2769.4	210.2	3031.4	21.6	2792.4	220.4	3002	0	3204.4	0	3205.2	47	3329.2	44.6	3237.2
		0.5	171.4	2984	216.2	3034.4	172	2963.4	214.2	3011.2	0	3205.2	0	3205.2	42.6	3338.2	62.2	3271
		0.99	240.8	3028.4	228.6	3036.2	247.4	3010.6	243.6	3019.6	16.4	3207.6	16.4	3207.6	95.8	3356.2	116	3286.8
	8	0.01	27	1298	33.2	1323.4	21.8	1317.4	28.2	1348.4	1	1529.6	1.4	1504.8	13.2	1481.8	18.6	1409.2
		0.5	45.8	1377.8	42.4	1366	34.8	1376.2	37.2	1376.6	12.6	1550.4	4.8	1528	20.4	1515.2	21.4	1419.4
		0.99	64.4	1385.4	64.8	1377.8	64	1389.4	55	1384	20.6	1554	23	1535.4	32.4	1519.4	29.8	1422.2

Chapter 4. Mixed Signals: An Empirical Study of the Alignment (and Misalignment) of Risk Signals with Actions and Outcomes in P2P Lending Markets

Introduction

Peer-to-peer (P2P) lending markets are internet-enabled platforms that create a venue for matching lenders and borrowers with lower overhead, and hence the potential for greater profit margins, than traditional financial institutions. These markets generally have three types of participants or players. *Borrowers* each request a loan amount and return period (e.g., 36-month or 60-month loans), and provide the details of their financial history and their personal information like their occupation and the purpose of the loan. The market owner or *platform* serves as a mediator between borrowers and investors, receiving the borrowers' requests and details and making potentially profitable loan requests available to lenders for selection. Though auction-based platforms (with pricing determined through a competitive process) have existed, we focus on markets in which the platform also prescribes an interest rate based on their analysis of credit information and the borrower's info. The platform in this case serves as an expert on risk assessment to prescribe the interest rate, rather than distributing this task among investors (as in the auction case). In many cases, the platform will also provide additional signals (beyond defining the interest-rate terms of the loan) such as a risk score, a grade label, and an estimated loss rate to facilitate differentiation of the loan applications for investors. The third set of players are *investors* who receive all of this information from the platform and decide how much to invest in each loan.

All players may behave strategically in this market. For example, although borrowers cannot change their financial history, they may be able to select the "purpose of the loan" strategically, given that the non-institutionalized and unsecured (i.e., with no collateral) market platform usually cannot verify this information. The platform, on the other hand, wants to select a format of information transmission (the loan terms, types of signals, and level of detail) that most efficiently clears loans and results in the largest profit over time. (Platforms are paid a small percentage of every payment from borrowers to the investors, aligning their incentives partially, but also receive a one-time fee.) Investors evaluate the ex post outcome of the loans precisely and predict the risk and benefit of each potential funding decision accurately.

Having collected a large dataset of publicly available loan information for over four years of loan origination requests (with all follow-up data through the completion of 36-month loan terms) from an anonymous lending platform, this study seeks to shed light on the interplay between the players in these markets, showing how signals from one class of participant effects the behavior of others using data analytics. In particular, we first explore the borrowers' disclosed personal information and analyze the response of the platform as well as the investors to these signals. Then, we analyze the response of investors to the platform signals and examine how closely investors follow the signals provided by the platform. Finally, we study the efficiency of both the borrower self-reported information as well as the P2P lending platform signals in predicting the success of each loan.

The paper is laid out as follows. Section 2 contains a brief review of selected literature on P2P lending markets related to our empirical study. Section 3 discusses the structure of P2P lending markets, as well as our dataset and empirical methodology. In the next section, statistical results are presented, with a discussion of our results, some conclusions, and directions for future research in the final section.

Background and literature review

In a P2P market, institutional or individual investors make unsecure loans to individual borrowers, meaning that borrowers do not provide any collateral to secure the loan. In these markets, information asymmetry between market players can affect the outcomes of the market (Lin, Prabhala & Viswanathan, 2013, Freedman & Jin, 2011, Miller, 2015, Serrano-Cinca & Gutiérrez-Nieto, 2016, Balyuk, 2016). Over the years, the P2P lending markets have implemented different mechanisms to address the information asymmetry issue and mitigate adverse selection among investors. In their early days, more P2P lending platforms used auction-like mechanisms in which borrowers posted their loan request, their maximum accepted interest rate, and a description of themselves. Investors would respond with offers at various interest rates resulting in the lowest offered interest rates winning the auction and funding the loan. Under this type of mechanism, the borrower self-descriptive information, the optional support of the loan application through the borrower's social network, as well the limited credit history information have each been considered in the literature in the context of adverse selection among investors (Lin et.al., 2013, Lin, Prabhala & Viswanathan, 2009, Collier & Hampshire, 2010, Emekter, Tu, Jirasakuldech & Lu, 2015, Dietrich & Wernli, 2016).

Indeed, the self-reporting information structure of borrowers and its effect on the P2P lending market has generated a great deal of research, mostly considering the effect of self-reported information on P2P lending market outcomes for markets in which borrowers could provide a short description about themselves and even post their pictures to attract investors and build trust. Herzenstein, Andrews, Dholakia, & Lyandres (2008) studied the effect of borrower personal characteristics like gender, race, and financial history on funding success in a traditional form of P2P lending market in which final loan interest rates were determined through an auction-based mechanism. Gonzales & Loureiro (2014) studied the effect of lender and borrower personal characteristics (perceived attractiveness, age, and gender) on outcomes of P2P lending markets and showed that loan success is sensitive to the relative age and attractiveness of lenders and borrowers. In 2015, they also reported, "certain borrower personal characteristics fuel interpersonal competition enough to impact lending decisions in suboptimal ways".

Additionally, some have focused on the effect of a self-reported loan purpose, considering the simultaneous effect of the "title" and "description of results" on loan applications. Mach, Carter & Slattery, (2014) showed that loans applications for small business purposes were twice as likely to have been funded than loans for other purposes. They also showed that controlling for application quality, loans for small businesses were charged almost one percentage point interest rate premium over non-business loans. In addition, Nowak, Ross & Yencha, (2018) analyzed the intriguing keywords in the borrower descriptions for small business loans and determined which keywords may result in a funded application for business purposes.

In the more recent generation of prominent P2P lending markets (which have gone away from auctions in favor of a posted rate, described next), borrower are now asked to self-report fewer pieces of information, likely an attempt to avoid the perception of race, gender, etc., having an effect on loan application outcomes. Currently, only a few pieces of information including occupation, length of occupation, income, and loan purpose are directly reported by borrowers and self-descriptive paragraphs about the loan applicant or loan purpose are no longer allowed. In addition, much of this information is randomly verified by the P2P lending platform.

Recent policies of prominent platforms provide more detailed information about borrower financial history to investors. They use a risk-based pricing structure to pre-define the interest rates for loan applications, sending signals about expected the risk and benefit of loan applications. Considering the financial history of each loan applicant as the basis to estimate the risk of each loan, the risk-based pricing can potentially decrease adverse-selection issues and increase efficiency in consumer loan markets. (See Walke, Fullerton, & Tokle, 2018; Cox, 2017; Adams, Einav & Levin, 2009; Edelberg, 2006; Xin, 2018). However, the accuracy of these platform signals in predicting loan outcomes has not yet been assessed in the depth of the current study. We find Serrano-Cinca, Gutiérrez-Nieto, López-Palacioz (2015) reported that on their studied data set (including 24,449 loans before 2011 from the LendingClub market), the "grade" signal was analyzed and found to be a significant indicator of delinquency in the market, but no further analysis of the veracity of platform signals was considered.

This evaluation of the existing body of literature illustrates that there is a need to provide a more holistic assessment of the role of all players in a P2P market and how their information revelation can alter other players' decisions, and thus change the market results. We have distinguished three main areas that need to be explored in this regard:

- How do borrowers use the "loan purpose" signal (i.e., the last remaining piece of the difficult-to-verify self-reported information) in current platforms?

- How do different investor decisions respond to these signals?

- To what extent do the borrower-reported loan purpose and the signals provided by the platform predict the risk and profitability of loan applications?

We investigate these questions in this research, while also taking the opportunity given by our dataset to compare the behavior of actual peer investors with the investment of institutional or high-net-worth investors on the same platform. We compare loans based on their risk and profitability as measured by delinquency ratio, internal rate of return, and return on investment.

The P2P Lending Market Place

In this paper, we study a P2P lending market data set that contains extensive information on both loan applications and loan outcomes. We specifically focus on the "*loan applications*", and the respective funded applications ("*loans*") initiated from 2011-01-01 until 2015-05-29, including those which were fully paid, charged off, or defaulted by 2018-05-29. (We only considered 36-month loans for this study as a matter of expediency.) This data set includes 284,983 loan applications of which 159,233 were funded (55.86%). The minimum loan amount in this period is $2000 and the maximum loan amount is $35000.

The P2P lending market of this study

In the market studied here, each borrower submits the requested loan amount and provides some information about her reason for borrowing the loan, annual income, occupation and employment status. Then, the platform pulls the borrower credit report from Experian, including information on her credit score, length of credit history, debt, number of accounts in the borrower's name, number of inquiries in the borrower's name, credit utilization ratios, balances utilization, and mortgage accounts. Based on the information provided by the borrower and her credit history, the platform decides about the posted interest rate that will be available to investors if they choose to fund this application. The platform also generates the "estimated loss rate", and assigns a "risk score" and a "grade score" to each loan application and provides these calculated and assigned scores as well as all information provided by borrowers available to investors.

Most current P2P lending markets have at least two funding channels, a "*fractional*" channel and a "*whole"* channel. In the fractional channel, any type of investor (institutional, high-net worth, or ordinary investors) can invest in any portion of a loan (usually above a predefined minimum like $25), and a loan will be originated only if at least a specific portion of a loan (for example, one platforms sets 70%) is funded[1]. In the "whole" channel, only institutional and validated high-net worth individuals can invest and they have to fund any selected loan in full amount. If an application is not funded in the whole channel, it can be moved to the fractional channel by the platform. In our data set, the whole channel itself initiated in April 2013. During our study, 213,759 loan applications were announced in the whole channel (75.01%) and 71,224 loan applications were announced in the fractional channel (24.99%). From 213,759 loan applications in the whole channel, 9,073 of them were not funded and then moved to the fractional channel (3.18%). Among these 9,073 moved loan applications, 5,633 applications are funded (62.08%), and 245 applications expired (2.7%).

When the minimum percent of a loan is funded, the P2P lending platform continues the verification process of borrower's information, which may result in cancellation of the loan application. In fact, if the data are not verifiable or are not accurate, or the risk of the loan is deemed higher than acceptable, the loan application will be cancelled (Balyuk, 2015). In addition, within the time window that the application is available in the market, the loan applicant can also withdraw her request. In our data set, 112,178 loan

[1] Borrowers have the option to be flexible and accept the loan if it is at least 70% funded, or to only accept the loan if it is fully funded.

applications were cancelled (39.35%), 9,445 loan applications were withdrawn (3.31%), and 4,127 loan applications were not funded within the available time window[2] and are expired (1.15%)[3].

There are two primary revenue sources for the P2P lending platform, a transaction fee ranging from 1-5% of loan amount (received from borrowers at the moment of loan origination) and service fees of 1% of each loan payment to investors (including any collection fees or recovery fees for late payments). In our data set, the average service fees for each loan is 2.17% of loan amount.

Dependent Variables:

Considering the loan application process at the P2P lending market, we categorize the dependent variables into three groups: P2P lending market (platform) signals, loan application outcomes, and ex-post loan outcomes. Table (1) lists these variables, with further detail provided below by category.

Platform Signals	
Posted Interest Rate	The interest rate assigned to each loan application by the market owner
Grade	The grade assigned at the time the loan application was created by the market owner and includes 7 levels AA, A, B, C, D, E , and HR which AA has the lowest risk and lowest interest rates.
Risk Score	A custom risk score that ranges from 1-11, with 11 being the best, or with the lowest risk.
Estimated Loss Rate	The Estimated Loss Rate is the estimated principal loss on charge-offs, and is provided by the market owner based on the historical data about similar loans.
Loan Application Outcome	
Funding Duration	The duration between origination of the loan application and the time loan is funded in hours.
Loan Funded Percentage	The percentage of the loan application amount funded by investors when the loan application is closed.
Ex-post Loan Outcome	
Delinquency	A binary variable which is equal to one if the loan became defaulted or charged off
Internal Rate of Return (IRR)	The actual interest rate at which the investment and payoff of the loan become equal
Return of Investment (ROI)	A financial metric calculated based on net profit of present values of loan payments divided by loan amount

Table (1) - Dependent Variables

P2P Lending Platform Signals:

The platform assigns four values to each new loan application: interest rate, grade, risk score, and estimated loss rate. The (posted) interest rate is the borrower's interest rate for the loan as decided by the platform (as opposed to a prevailing market rate). The "grade" in most P2P lending markets is a label

[2] In the fractional channel, the available time window is 14 days.
[3] From 112,178 cancelled applications, 24,990 loan applications are completed but and later did not received any loan ID. We considered this group as cancelled loan applications.

used to partition the loan applications based on their total risk and benefit. However, it is possible that a loan with a better grade receives a higher interest rate. The “risk score” is an ordinal number between one and eleven, with higher values representing less risky loan applications. Finally, the “loss rate” that is being estimated is defined as the percentage of unpaid principal (in a reporting time period) deemed “uncollectible”, i.e., taken as a loss.

Loan Application Outcomes (investor response)

We consider two metrics to measure the response of lenders to loan applications: *percentage of loan funded* and *funding duration.* The percentage of loan funded is the fraction of the loan application amount that is funded within the available time window for loans in the fractional channel. (This is 100% for the whole channel by definition, so will only be discussed in the context of the fractional channel.) The funding duration is another metric to measure the attractiveness of a loan application to investors, the time between announcement of the loan application by the platform and the time the loan is funded. This time varies for different loan applications based on their risk and benefits, and varies dramatically across the two channels. All analysis on these variable considers only completed and expired loan applications[4] in order to focus on investor decisions as opposed platform decisions.

Measurement of Ex-post Loan Outcomes:

In general, there are two approaches to compare the ex post outcomes of loans in consumer loan markets. In the first approach, the risk refers to the possibility of delinquency for each loan (Wiginton 1980, West, 2000, Baesens, Gestel, Viaene, Stepanova, Suykens, Vanthienen, 2003), while in the second approach, the risk is defined based on the *profitability of each loan* (Thomas, 2000, and Serrano-Cinca & Gutiérrez-Nieto, 2016). The possibility of delinquency is the core concept in defining a credit score and differentiating customers in the credit market. In fact, due to the lack of precise data on customer

[4] Those funded loan applications that are completed but their loan ID were not registered are also considered since cancellation happened after selection by investors as opposed to preemptive cancellation by the platform.

products, calculation of the loan profitability can be challenging for banking systems[5] (Baesens et. al, 2003). Therefore, credit scores and interest rates in the loan markets are mainly defined based on possibility of delinquency. However, there is an opportunity in P2P lending markets to help investors to attain a more precise understanding of the risks of each loan in the market. In large P2P markets, loan payments and the charge-off sale of loans are available to investors in order to help them to mitigate adverse selection and estimate benefit of each loan application.

Profitability Measurement

There are two well-known financial formulas to evaluate the efficiency of each investment in the consumer market, Internal Interest Rate (IRR) and Return on Investment (ROI). IRR is the actual interest rate at which the investment and net present value (NPV) of a loan become equal (Gallo 2016). IRR has been used in recent research as a measure of profitability in the P2P lending markets, for example by Serrano-Cinca and Gutiérrez-Nieto (2016) who provide a decision support system for P2P lending markets that relies on an estimation of IRR for each loan. They illustrate IRR's superiority over default probability calculate the *effective* loan interest rates of each investment. While IRR can indeed be helpful for ranking loans with good ex post outcomes (when the total return exceeds the initial investment, whether full paid or not) we find that IRR is less informative when charged-off and defaulted loans are considered.

Indeed, using the standard IRR method, one assumes that the funds released from a loan are reinvested in another loan yielding IRR equal to the previous loan (Galo, 2016). However, this assumption is not always possible in loan markets. In particular, for delinquent and charged off loans, this assumption is completely misguiding in measuring the loan profits and ranking loan outcomes. For example, consider loans as in Table (2). In this example, there are four $1000 loans with different

[5] "Customers may own different products ranging from mortgages to credit cards, and may use different channels, ranging from bank branches to online banking. All of these combined factors make it difficult to obtain precise data on customer profitability, and researchers complain about the lack of enough data to investigate profit scoring" (Serrano-Cinca & Gutiérrez-Nieto, 2016)

payments. The return of the first two loans are smaller than the loan principle; however, the first one retuned \$800 in one month, while the second loan returned the same amount in four months. We prefer the first loan outcomes to the second loan since the returned money from the first loan can be invested sooner on another (hopefully profitable) investment; however, IRR shows a less negative interest rate for the second loan, prioritizing it over the first one. The reason is IRR assumes that the investment of the returned money will be repeated on a loan with *the same interest rate*, and losing money will continue with the same pattern.

	Loan 1	Loan 2	Loan 3	Loan 4
Initial Investment	-1000	-1000	-1000	-1000
Payment month 1	800	200	300	1200
Payment month 2	0	200	300	0
Payment month 3	0	200	300	0
Payment month 4	0	200	300	0
IRR (monthly)	-20.00%	-8.36%	7.71%	20.00%

Table (2) - Comparison of loan profitability based on IRR for four examples

The third and four loans of this example, shown in table (2), discuss two other loans that their returned money is above the initial investment. This time the shorter repayment has higher IRR, which seems appropriate. However, the annual IRR for the fourth loan is 240%, which means that the investment is more than doubled in one year. Again, the assumption of IRR for reinvesting the received payments in loans with the same interest rate provides misleading results.

As an alternative, ROI calculates the profit to cost ratio of an investment, with the NPV of future cash flows (based on a particular discount rate) compared to the original investment. In a P2P lending market, ROI calculates the ratio of summation of present value (from the perspective at the moment of origination) of all loan payments minus the (undiscounted) loan amount, to the loan amount. To calculate the present value of each payment, we define the discount rate to be the average monthly inflation rate within the loan origination date and each payment date. Table (3) represents the NPV and ROI of four discussed examples in table (2) and shows that ROI can clearly rank the ex post outcomes of these loans (in particular the negative outcomes) better than IRR.

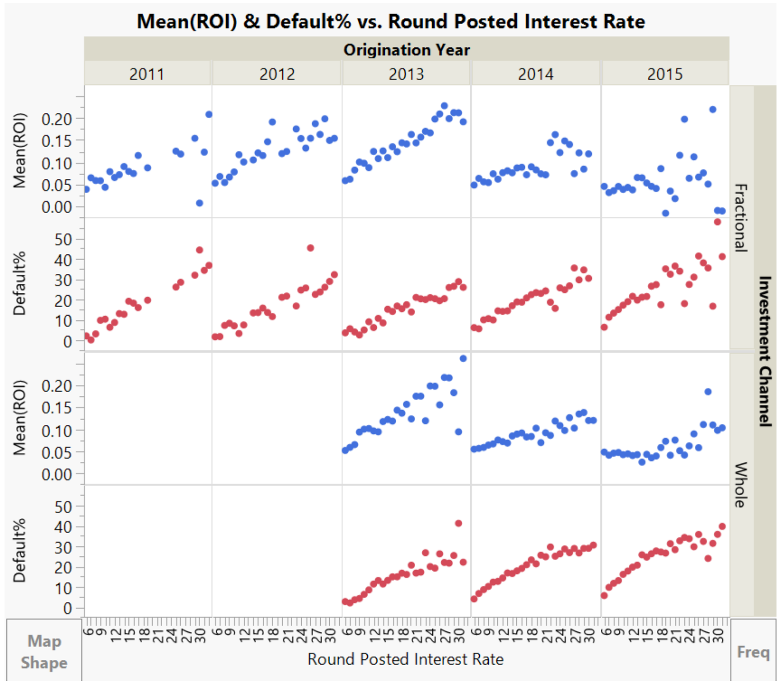
	Inflation rate	Average inflation rate until month t	PV Loan 1	PV Loan 2	PV Loan 3	PV Loan 4
Initial Investment	2.50%	2.50%	-1000	-1000	-1000	-1000
Payment month 1	2.70%	2.60%	779.7271	194.9318	292.3977	1169.591
Payment month 2	2.40%	2.53%	0	190.2391	285.3587	0
Payment month 3	2.20%	2.45%	0	185.9919	278.9879	0
Payment month 4	1.90%	2.34%	0	182.3259	273.4888	0
	NPV		-220.273	-246.511	130.2331	169.5906
	ROI		-22.03%	-24.65%	13.02%	16.96%

Table (3) - NPV and ROI for examples introduced in table (2)

Figure (1) shows the average of ex post outcomes of loan applications within years in both the fractional and whole channels. In this figure, we observe that for loans funded in the whole channel, as the posted interest rate increases the average ROI of loans also increases. However, the slope of this increase has decreased from 2013 to 2015. In particular, in 2015, for the low and medium values of the posted interest rate, ROI is almost constant. However, for the higher values of the posted interest rate, the average ROI increases as the posted interest rate increase.

In addition, while results of the whole channel show a consistent pattern of increasing ROI as the posted interest rate increases, investors in the fractional channel could not consistently improve their profit by selecting higher interest rate loans, and in particular in 2015 their outcomes do not follow the increasing trend of all other sub-figures.

Figure (1) – Average loan outputs in the fractional and whole channels within years



A more detailed descriptive analysis of outcomes for the fractional and whole channels using different explanatory variables is shown in the Appendix.

Data Description

Table (4) lists the remaining independent variables considered in this paper categorized in three groups: loan specifications, borrower specifications, and economic conditions. Loan specification variables include the loan amount, the loan purpose, and the loan origination year. Borrower specifications are variables about financial and credit history of borrowers. Economic conditions include the inflation rate and Federal Reserve rate at each loan origination date.

Descriptive Statistics of P2P Lending Market

In this research, we use a data set of P2P lending market that contains extensive information about individual loan applications. We categorize these variables to three groups. The first group discusses the borrower's personal characteristics including income range and homeownership. The second group describes the borrower's credit history and contains variables such as credit score range, number of delinquencies over 30 days, delinquencies within the last 7 years, inquiries within the last six months, and revolving balance. The third group refers to loan application conditions and contains the loan application purpose, the loan application amount, the loan application date, as well as the inflation rate and Federal Reserve rate in the month of loan application submission. Table (4) lists definitions of these variables.

Loan Specifications	
Loan Application Amount	The dollar amount of the loan application.
Loan Purpose	12 loan purposes including Auto/Motorcycle/RV/Boat, Baby and Adoption, Business, Debt Consolidation, Green Loans, Household Expenses, Medical and Dental, Not Available, Other, Taxes, Vacation or Special Occasion, and Wedding Loans
Loan Origination Year	The year the loan application is originated online on the website
Inflation Rate (monthly)	Monthly US inflation rate (https://www.usinflationcalculator.com/inflation/current-inflation-rates/)
Federal Reserve Rate	Monthly US Federal Funds Rate (https://www.federalreserve.gov/monetarypolicy/openmarket.htm)
Borrower Personal Characteristics	
Income Range Description	Five ranges of income including, $0, $1-$24,999, $25,000-$49,999, $50,000-$74,999,$75,000-$99,999, and $100,000+.
Homeownership	A binary variable which is equal to one if the borrower has a mortgage on her credit profile or provides documentation confirming she is a homeowner
Borrower Credit History	
Credit history Length	The duration between opening the borrower's first reported credit card and the date of the loan application origination measured in decimals in years.
Bankcard Utilization	The percentage of available revolving credit that is utilized.
Debt to income with the loan	the ratio of the monthly debt including the proposed loan application to monthly income measured in decimals
delinquencies within last7 years	Number of delinquencies in the past 7 years
delinquencies over 30 days	Number of delinquencies for more than 30 days
Inquiries last 6 months	Number of inquiries in the past six months before loan application
total inquiries	Total number of inquiries before loan application
total open revolving accounts	Number of open revolving accounts at the time of the loan application
revolving balance	Amount of revolving credit at the time of the loan application in dollars
revolving_available_percent	Percentage of revolving credit available in decimal
Fico Score Range	A measure of consumer credit risk, based on FICO credit reports in 14 ranges NA, <600, 600- 619, 620-639, 640-659, 660-679, 680-699, 700-719, 720-739, 740-759, 760-779, 780-799, 800-819, 820-850
Credit Score Range	A measure of consumer credit risk, based on bureau credit reports in 12 ranges NA, <600, 600- 619, 620-639, 640-649, 650-664, 665-689,690-701, 702-723,724-747, 748-777,778+

Table (4) - List of Variables

Empirical Modeling

In this section, we address three questions about the three player types in a P2P lending market. First, we explore how borrowers can adjust their self-reported information to receive lower interest rates. We also examine if their self-reported information can influence investors and attract more investment in a shorter amount of time. Secondly, we study how platform signals influence investors' decisions. Thirdly, we compare loans based on their ex post outcomes measured by delinquency and profitability, and analyze how the borrowers' self-reported information, personal characteristics, and credit history as well as the platform signals can significantly predict the outcome of loans in the market. In addition, we examine the effect of time on accuracy of the platform signals in predicting loans' ex post outcomes.

1- Evaluation of Effect of the Borrowers' Self-Reported Information

In today's prominent P2P lending markets, most borrower information consists of verified details about their credit history. Only a small portion of borrower information is self-reported, and the platform randomly checks pieces of this information such as income range, occupation, employment status, and length of employment. The only fully unverified piece of information borrowers can provide to mislead the platform and/or investors is the *loan purpose*. Therefore, we first explore the sensitivity of platform and investor decisions to the stated loan purpose, i.e., we assess the impact of loan purpose on the loan interest rate, funding duration, and percentage of loan funded.

Effect on P2P Lending Platform: Defining Posted Interest Rate

In the first model, we assess the effect of loan purpose on the platform's posted interest rate for a loan by estimating a regression model as given by model (1).

$$PstRate_i = \alpha + \beta \times LnPrps_i + \mu \times LnSpc_i + \gamma \times BrrPr_i + \delta \times BrrCr_i + \epsilon_i \quad (1)$$

The dependent variable is the posted interest rate for the loan application ($PstRate$), and the independent variable of primary interest is the loan purpose ($LnPrps_i$), while we control for other loan specifications (vector $LnSpc_i$), the borrower personal characteristics (vector $BrrPr_i$), and the borrower credit history factors (vector $BrrCr_i$), as described in table (4). After checking our model for heteroscedasticity, we found that the variances of the residuals are correlated with our explanatory variables, requiring the models to be rerun with a relaxation of the homoscedasticity assumption. Model (1) therefore uses a robust GLS (Generalized Least Square) model which adjust the observation with the variance to estimate the coefficients under these conditions.

Table (5) illustrates part of the results of this model that explains the relationship between the loan purpose and the posted interest rate. This result shows that the adjusted R^2 is 0.61, and seven out of the eleven loan purpose categories have a statistically significant relationship with the posted interest rate. The model assumes a value of zero for "Taxes" and the sign and value of coefficient of other categories are explained comparing to this category. Accordingly, sign of "Auto/Motorcycle/ RV/Boat" and "Debt Consolidation" are negative compared to "Taxes", implying that loan applications in these categories would expect to receive a lower interest rate compare to their peers in the "Taxes" category. In contrast, the coefficients of "Baby & Adoption", "Green Loans", "Business", "Medical/Dental" and "Wedding Loan" categories have positive signs compared to "Taxes" and they would expect to receive a higher interest rate. Comparing values reveals that, among all groups, "Debt Consolidation" and "Auto/Motorcycle/RV/Boat" seem the best categories to receive the lowest posted interest rates, and "Baby & Adoption", "Green Loans" and "Business" receive the highest interest rates, while "Other" and "Not Available" categories have no statistically significant influence on the posted interest rate. The full results of this model can be found in Appendix, table A.4.

Effect on investors: Percentage of loan-funded

In the second and third models, we study whether or not stating specific loan purposes can influence investors. In model (2), which is defined only for the loan applications in the fractional channel, we examine how the loan purpose can affect the amount of investment.

$$FndPrc_i = \alpha + \beta \times LnPrps_i + \mu \times LnSpc_i + \gamma \times BrrPr_i + \delta \times BrrCr_i + \varphi \times MrkSgn_i + \epsilon_i \quad (2)$$

The dependent variable is the percentage of loan funded ($FndPrc$), and the independent variable of primary interest is the loan purpose ($LnPrps_i$). We control for loan specifications (vector $LnSpc_i$), the borrower personal characteristics (vector $BrrPr_i$), the borrower credit history factors (vector $BrrCr_i$), as well as the platform signals (vector Mrk_Sgn_i) including the posted interest rate, risk score, grade score, and estimated loss rate.

In this analysis, we have focused on respond of investors to stated loan purpose. Accordingly, we only considered the completed or expired loan applications in the market, and excluded "Withdrawn" or "Cancelled" loan applications. Similar to model (1), we have found that the model has the heteroscedasticity so we have used a weighted regression model to address this issue. Furthermore, because of the high correlation between two important explanatory factors of the platform signals, the posted interest rate and the estimated loss rate, we used the log transformation of the estimated loss in all our next models.

Table (5) represents part of the results of this model that explains the relationship between the loan purpose and the percentage of loan funded. Though R^2 is relatively low for this model (0.19), our findings show that four categories out of eleven categories of loan purposes have a statistically significant relationship with percentage of loan funded. In particular, the loan applications that state "Not Available" as their purpose would expect 0.197 lower percentage of loan funded compared to "Taxes". Moreover, "Business", "Household Expenses", and "Other" categories attract statistically significantly lower amount of investment in comparison to "Taxes", while other stated loan purposes do not have any statistically significant influence on percentage of loan funded. The full results of this model can be found in Appendix, table A.4.

Effect on investors: Funding duration

The third model is very similar to model (2) and assesses the relationship between the loan purpose and the funding duration ($FnDur$) both in the fractional and whole channels. However, since the funding duration values range over several orders of magnitude, in this model we use the *log transformation* of the funding duration as the dependent variable. In addition, to address the heteroscedasticity, we also have employed a GLS model.

$$Log(FnDur_i) = \alpha + \beta \times LnPrps_i + \mu \times LnSpc_i + \gamma \times BrrPr_i + \delta \times BrrCr_i + \varphi \times MrkSgn_i + \epsilon_i \quad (3)$$

Table (5) illustrates the results of this model that explain the relationship between loan purpose and the funding duration, for fractional and whole channels, separately. Investors in both channels have a significantly different response to the various loan purposes. In the fractional channel, which R^2 is 0.34, eight of the eleven categories have statistically significant relationship with the funding duration. Investors in the fractional channel are sensitive and responsive to the loan purpose and “Not Available”, “Business”, “Other”, and “Household Expenses” categories have the longest funding duration among significant loan categories in this channel. This result is in accordance with the results of model (2) in which loan applications in “Not Available”, “Business”, “Household Expenses”, and “Other” categories would expect lower percentage of loan funded in the fractional channel.

	Posted Interest Rate (Model 1)			Loan Funded Percentage (Model 2)			Loan Application Duration (Model 3)					
							Fractional			Whole		
Loan Purpose	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**
Business	0.0051	***	0.0010	-0.0428	***	0.0084	27.0759	***	4.0523	9.5345	***	0.6069
Household Expenses	0.0001		0.0010	-0.0233	**	0.0081	16.3868	***	3.8913	7.5323	***	0.5583
Debt Consolidation	-0.0084	***	0.0009	0.0000		0.0079	9.3989	*	3.7745	6.3240	***	0.5401
Auto/Motorcycle/RV/Boat	-0.0041	***	0.0011	-0.0023		0.0090	1.9630		4.3570	5.9846	***	0.6892
Wedding Loans	0.0037	**	0.0013	-0.0042		0.0111	9.2024	*	5.3396	5.3179	***	0.8694
Other	-0.0008		0.0010	-0.0203	*	0.0082	21.6993	***	3.9333	3.5427	***	0.5652
Vacation/Special Occasion	-0.0007		0.0013	-0.0119		0.0111	14.2633	**	5.3408	2.0617	**	0.7785
Medical / Dental	0.0050	***	0.0011	-0.0151		0.0096	12.1102	**	4.5847	1.1437	*	0.6316
Baby & Adoption	0.0088	***	0.0023	-0.0395		0.0202	10.2444		9.6191	-0.3646		2.1255
Not Available	-0.0036		0.0033	-0.1969	***	0.0355	96.9224	***	16.9003	-7.0220	**	2.1495
Green Loans	0.0078	*	0.0034	0.0241		0.0330	22.2644		16.3982	-7.7017	**	2.8331
Platform Signals				**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**
Posted Interest Rate				1.7985	***	0.0903	-164.1561	***	40.6011	68.6300	***	7.2960
Estimated Loss Rate				-0.2992	***	0.0096	59.1275	***	3.4963	-1.3355	**	0.4817
Grade												
Grade A				0.1537	***	0.0081	-50.5552	***	2.6658	3.6024	***	0.2694
Grade B				0.2540	***	0.0114	-97.0026	***	4.0191	1.4334	***	0.3776
Grade C				0.2937	***	0.0139	-118.0184	***	5.1043	2.1000	***	0.4807
Grade D				0.3041	***	0.0164	-116.6755	***	6.2843	0.3033		0.6262
Grade E				0.2997	***	0.0191	-115.7482	***	7.6854	1.8195	*	0.8245
Grade HR				0.1804	***	0.0206	-56.5830	***	8.3524	1.7386	*	1.0402
Risk Score												
Risk Score = 2				0.1050	***	0.0062	-9.2909	**	3.5429	14.5190		20972.2500
Risk Score = 3				0.0948	***	0.0061	1.8889		3.5028	15.7422		20972.2500
Risk Score = 4				0.1027	***	0.0060	6.4147	*	3.4298	19.7285		20972.2500
Risk Score = 5				0.1198	***	0.0062	1.9532		3.5018	26.8145		20972.2500
Risk Score = 6				0.0992	***	0.0064	14.1818	***	3.5883	22.7079		20972.2500
Risk Score = 7				0.1026	***	0.0068	18.5230	***	3.7129	23.1178		20972.2500
Risk Score = 8				0.1052	***	0.0072	33.4623	***	3.7912	23.5758		20972.2500
Risk Score = 9				0.1104	***	0.0083	31.2250	***	4.0547	25.6408		20972.2500
Risk Score = 10				0.0959	***	0.0098	17.1894	***	4.3750	25.7591		20972.2500
Risk Score = 11				0.0509	***	0.0112	27.5615	***	4.6126	35.1369		20972.2500
Number of Observations	284,983			50,709			50,709			137,641		
R-Square Adjusted	0.6109			0.1912			0.3435			0.3585		

*** indicates p-value<0.001, ** indicates p-value<0.01, * indicates p-value <0.1

Table (5) – Part of the Results of Model (1), (2), and (3)

In the whole channel, R^2 is 0.36 and ten out of the eleven categories have statistically significant relationship with the loan purpose. Similar to the fractional channel, loan applications that stated “Others”, “Household Expenses” and “Business” as their loan purpose would expect a longer funding duration. However, loan applications in “Not Available” and “Green Loans” categories would expect on average a shorter funding duration in the whole channel. It is noteworthy that despite the relatively lower posted interest rates for “Auto/Motorcycle/RV/Boat” and “Debt Consolidation”, investors in the whole channel are significantly less interested in these categories, and loans in these categories have the third and fourth longest funding durations. However, despite the high posted interest rates for the “Green Loans” which is expected to represent the high risk of these loans, in the whole channel, loan applications in this category are the most popular loans based on the funding duration. The full results of this model can be found in the Appendix, table A.4.

Taking all of these models together, we found that when an investor states “Business” as the loan purpose, she would expect a higher posted interest rate, lower investment, and longer funding duration. Moreover, if she states “Others”, she would expect no statistically significant change in the posted interest rate, a little lower investment, and a little longer funding duration. In addition, if she states “Debt Consolidation” as the loan purpose, she would expect a lower interest rate and a mildly longer funding duration.

2- Evaluation of Effect of the Platform Signals on Investors

In this section, we answer to the second question of this paper and analyze how investors respond to the platform signals. In fact, the main role of the P2P lending platform is to provide appropriate signals about the loan applications to mitigate the adverse selection among investors. In the original model of P2P lending mechanisms, the platform’s only role was to assign signals like grade and risk score to the loan applications, and investor responses to borrower maximum acceptable interest rate were the main factor to define the final interest rates. However, in today’s prominent P2P lending mechanisms, the platform is the key player who defines and posts the interest rates, and investors have no direct influence in defining interest rates.

In recent P2P lending markets, investors have no direct role in defining the posted interest rate; however, the platform rigorously needs to consider the response of investors to its signals and adjust them based on their expected response. Although non-expired loan applications can show the investors' satisfaction with the posted interest rate and the platform signals, more sensitivity analyses can be helpful for the platform to adjust its signals based on the investors' expectations. In fact, an expired loan application is a *lost opportunity* for the platform, and the platform will clearly seek to minimize these lost opportunities and increase the investors' eagerness to participate in the market. Clearly, a long "funding duration" or a low "percentage of loan funded" are a matter of revealed investor preferences verifying the influence of the platform signals on the behavior of investors. Therefore, in this section, we analyze the relationship between the P2P platform signals and the funding percentage (of the total requested loan amount in the fractional channel) and the funding duration in both the fractional and whole channels.

Effect on investors: Percentage of loan funded

The model for this analysis is the same as model (2) that specifically explores the loan applications offered in the fractional channel. For this analysis, the main focused independent variables are the platform signals and we control for the loan purpose, loan specifications, borrower personal characteristics, and borrower credit history to analyze the relationship between the P2P platform signals and the percentage of loan funded. We again refer to table (5) that illustrates the results of model (2) and this time focus on the platform signals: posted interest rate, estimated loss rate, grade, and risk score.

The results in table (5) show that all platform signals have statistically significant influence on the percentage of loan funded and investors in the fractional channel respond to the platform signals about the loan applications. The sign of the "Posted Interest Rate" (positive) and the "Estimated Loss Rate" (negative) show that increasing the posted interest rate and decreasing the estimated loss rate would result in higher percentage of loan funded.

In model (2), in which we use the grade "AA" as the reference with effect equals to zero, the sign of the level coefficients compared to the level "AA" are all positive. Indeed, worsening the grade of a loan application would consistently increase the percentage of loan funded. However, if the loan labeled as grade "HR" (i.e., high-risk), the borrower would expect a lower percentage of loan funded that grades "B", "C", "D", and "E".

In addition, all levels of the risk score have a statistically significant relationship with the funded percent of the loan amount. However, the trend of the risk score coefficients in the model is not consistent and has very limited variations, except for the highest and lowest risk scores. Indeed, a loan application with a risk score of one or eleven would expect a smaller percentage of loan funded comparing to other risk scores.

Effect on investors: Funding Duration

In this section, we explore the investor response to the platform signals and analyze the relationship between the platform signals and the funding duration both in the fractional and whole channels. The model for this analysis is the same as model (3) and its results are presented in table (5). In this analysis, the platform signal is the independent variable of primary interest and we control for the loan purpose, loan specifications, borrower personal characteristics, and borrower credit history.

The analysis for the fractional channel reveals that the posted interest rate is inversely related to the funding duration and increasing the posted interest rate would shorten the funding duration. This result is in accordance with our previous findings about the influence of increasing the posted interest rate on investors' willingness higher investment in the fractional channel. Moreover, as expected, increasing the estimated loss rate increases the funding duration in the fractional channel. It is also observable that all grade levels have a statistically significant relationship with the funding duration. Among all grades, loan application labeled as grade "C", "D", and "E" would expect the shortest funding duration while "AA", "HR", and "A" would expect the longest funding duration in the fractional channel. This U pattern is also similar with our previous funding about the percentage of loan funded. Finally, eight levels out of ten grade levels have statistically significant relationship with the funding duration. In general, increasing the risk score (toward better loan applications), with slight variation, would increase the funding duration in the fractional channel.

The results of the whole channel however show a dissimilar relationship between platform signals and funding duration. First, the posted interest rate and the funding duration in the whole channel are directly related and increasing the posted interest rate would increase the funding duration. Secondly and more surprisingly, the estimated loss rate has the inverse relationship with funding duration and increasing the estimated loss rate would result in a decrease in funding duration in the whole channel. Thirdly, five out of the six grade levels have statistically significant relationship with the funding duration. Loan applications with grade "AA" would expect the shortest funding duration in the whole channel. However, the funding duration increases for the grade "A" and then is cyclic pattern is repeated as the grade is decreasing. Finally, the last platform signal, the risk score, has no statistically significant relationship with funding duration and investors in the whole channel seem to be completely ignoring this signal.

Taking all of these modes together, in the fractional channel, investors are eager to invest in riskier loan applications. More loan applications with a higher posted interest rate and lower risk score (toward riskier application) would surprisingly increase investor activity in this channel in general, indicating risk-seeking behavior, with the caveat that their interest in "HR" applications is similar to "A" grade. Investors in the whole channel, on the other hand, are in general risk averse and more interested in choosing applications with a lower posted interest rate, which is reasonable since they have to invest in the whole loan amount. However, their response to other platform signals, which describe the applications risks, show that they do not follow signals and their decisions are not aligned with the natural interpretation of the platform messages.

3- Evaluation of loan ex post outcomes

In this section, we address the third question of this paper and analyze the ex post outcomes of loans. We calculate three metrics to measure the loan outcomes, the possibility of delinquency, the internal rate of return (IRR), and the return of investment (ROI). First, we explore the relationship between the borrower signals and each loan's ex post outcomes to measure the reliability of the self-reported loan purpose. Then, we investigate the relationship between the platform signals and the loan ex post outcomes, and study whether each of these signals are significant predictors of the loan application risk and profitability.

Evaluation of Outcomes of Loan Applications Considering the Borrower Signals

In this section, we analyze whether or not the ex post outcomes of loans show significant differences between similar loans with different stated purposes. The dependent variables in model (4), (5), and (6) are delinquency ($Delinq_i$), internal rate of return (IRR_i), and return on investment of loans (ROI_i), respectively. In this section, the independent variable of primary interest is the loan purpose while we control for loan terms, borrower personal characteristics, borrower credit history, and the platform signals, as defined in tables (1) and (4).

Delinquency

In model (4), our dependent variable is a binary outcome: delinquent loan and not delinquent loan. Accordingly, in the model, we set the delinquency variable ($Delinq_i$) equal to one if the loan was defaulted or charged off. To assess the relationship between the loan purpose and the possibility of delinquency, we compared the results of the probit, logit and extreme value models for the default binary variable, and found that the logit model provides a more robust result considering the maximum likelihood and log likelihood. Therefore, we used a binary logit model to explore the relationship between the loan purpose and the possibility of delinquency.

$$Delinq_i = \alpha + \beta \times LnPrps_i + \mu \times LnSpc_i + \gamma \times BrrPr_i + \delta \times BrrCr_i + \varphi \times MrkSgn_i + \epsilon_i \quad (4)$$

Our independent variable of primary interest in this analysis is the loan purpose, while we control for other loan specifications, the borrower personal characteristics, the borrower credit history factors, and the platform signals as described in tables (1) and (3). Table (6) represents the results of this model.

The results reveal that seven out of eleven categories, “Business”, “Baby & Adoption”, “Vacation/Special Location”, “Other”, “Medical/ Dental”, “Household Expenses”, and “Not Available” all with positive coefficients compared to “Taxes”, have statistically significant relationship with “default” outcome. Among these seven significant loan purposes, “Business” has the largest positive coefficient in the model which indicates that if the borrower states “Business” as her loan purpose, the odds of getting default would increase compared to other groups of loan purpose.

Internal Rate of Return (IRR)

In model (5), our dependent variable is the internal rate of return (IRR_i) and we consider the same sets of independent variable as model (4). For this model, we also have released the homoscedasticity assumption by developing a weighted regression model in order to achieve more robust findings.

$$IRR_i = \alpha + \beta \times LnPrps_i + \mu \times LnSpc_i + \gamma \times BrrPr_i + \delta \times BrrCr_i + \varphi \times MrkSgn_i + \epsilon_i \quad (5)$$

Table (6) represents the results of this model. Our findings show that the adjusted R^2 for this model is 0.61 and only three categories out of the eleven loan purpose categories have statistically significant relationship with IRR of the loans. Among these three categories, “Auto/Motorcycle/RV/Boat” has the highest IRR, and the “Business” category has the lowest IRR.

Return on Investment (ROI)

Now we turn our attention to a more reliable metric for measuring the outcomes of loans in the P2P lending market, the return on investment that is the dependent variable in model (6). For this model, we found that the variances of the residuals are correlated with our explanatory variables so we have released the homoscedasticity assumption and developed a GLS (Generalized Least Square) model with inverse variance as weights to estimate the coefficients under these conditions

$$ROI_i = \alpha + \beta \times LnPrps_i + \mu \times LnSpc_i + \gamma \times BrrPr_i + \delta \times BrrCr_i + \varphi \times MrkSgn_i + \epsilon_i \ (6)$$

Our results in table (6) reveal that ROI of loan applications is almost not related to the loan purpose and ten out of eleven loan purpose categories do not have statistically significant relation with the ROI of loans. However, the only statistically significant loan purpose is the "Other" category that has the *positive* relationship with profitability of loans. Therefore, it means that considering the posted interest rates, loan purposes by themselves cannot predict any difference on loan outcomes and any extra attention to interpretation of loan purpose can be misguiding.

Evaluation of Outcomes of Loan Applications Based on the P2P Platform Signals

In this section, we evaluate how the platform signals are statistically significant indicators to predict the ex post outcomes of loan applications. We study the relationship between ex post outcomes of loan applications based on delinquency, IRR, and ROI, and the market signals including the posted interest rate, the estimated loss rate, the grade, and the risk score.

The models for these analyses are the same as models (4), (5), and (6). The dependent variables in model (4), (5), and (6) are delinquency ($Delinq_i$), internal rate of return (IRR_i), and return on investment of loans (ROI_i), respectively. In these models, the independent variables of primary interest are the market signals, while we control for the loan purpose, the loan specifications, the borrower personal characteristics, and the borrower credit history as defined in tables (1) and (4).

Delinquency

The model for this analysis is the same as model (4) and again we refer to table (6) for its results. As this table illustartes the posted interest rate, the estimated loss rate and the risk score have statistically significant relationship with the possibility of delinquency. The sign of coefficients of these variables in the model follow expectation. In particular, the possibility of default would increase as the posted interest rate and the estimated loss rate increase. In addition, the possibility of delinquency smoothly and consistently decreases as the risk score changes from 1 to 11 (toward less risky applications).

However, the analysis shows no statistically significant relationship between the loan grade and the delinquency rate, and grades are completely unsuccessful in predicting the delinquency rate of loan applications.

	Default (Model 4)			IRR (Model 5)			ROI (Model 6)		
Loan Purpose	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**
Auto/Motorcycle/RV/Boat	-0.0781		0.1029	-0.0008	*	0.0005	-0.0035		0.0031
Baby & Adoption	0.4104	*	0.2038	-0.0006		0.0011	-0.0082		0.0072
Business	0.5306	***	0.0913	-0.0010	*	0.0004	0.0029		0.0028
Debt Consolidation	0.1333		0.0848	-0.0002		0.0004	-0.0006		0.0025
Green Loans	0.2814		0.3254	-0.0009		0.0018	-0.0109		0.0117
Household Expenses	0.2050	*	0.0875	-0.0005		0.0004	0.0031		0.0026
Medical / Dental	0.2099	*	0.0971	0.0002		0.0005	0.0043		0.0030
Not Available	0.1476	*	0.2720	-0.0004		0.0014	0.0052		0.0093
Other	0.2107	*	0.0876	-0.0009	*	0.0004	0.0064	*	0.0026
Vacation/Special Occasion	0.2171	*	0.1156	-0.0004		0.0006	0.0000		0.0037
Wedding Loans	-0.1939		0.1227	-0.0003		0.0005	-0.0032		0.0036
Platform Signals									
Posted Interest Rate	2.4334	**	0.9089	-0.0351	***	0.0045	1.6143	***	0.0305
Estimated Loss Rate	0.2279	**	0.0855	0.0082	***	0.0004	-0.0124	***	0.0028
Grade									
GRADE="A"	0.0243		0.0534	-0.0016	***	0.0003	0.0027		0.0018
GRADE="B"	0.0242		0.0736	-0.0034	***	0.0004	0.0046	*	0.0025
GRADE="C"	0.0486		0.0894	-0.0023	***	0.0005	0.0048		0.0031
GRADE="D"	0.0398		0.1053	-0.0009		0.0006	0.0028		0.0037
GRADE="E"	0.0628		0.1236	0.0027	***	0.0007	0.0035		0.0044
GRADE="HR"	0.0378		0.1381	0.0040	***	0.0007	0.0137	**	0.0048
Risk Score									
RISK_SCORE=2	-1.0870	***	0.2214	0.0055	***	0.0014	-0.1159	***	0.0082
RISK_SCORE=3	-1.1247	***	0.2212	0.0051	***	0.0014	-0.1183	***	0.0082
RISK_SCORE=4	-1.1326	***	0.2208	0.0044	**	0.0013	-0.1166	***	0.0082
RISK_SCORE=5	-1.1887	***	0.2212	0.0045	***	0.0014	-0.1184	***	0.0082
RISK_SCORE=6	-1.2278	***	0.2216	0.0037	**	0.0014	-0.1141	***	0.0082
RISK_SCORE=7	-1.2752	***	0.2221	0.0035	**	0.0014	-0.1169	***	0.0083
RISK_SCORE=8	-1.3689	***	0.2225	0.0043	**	0.0014	-0.1173	***	0.0083
RISK_SCORE=9	-1.5018	***	0.2240	0.0019		0.0014	-0.1173	***	0.0083
RISK_SCORE=10	-1.6273	***	0.2273	0.0019		0.0014	-0.1167	***	0.0084
RISK_SCORE=11	-1.8287	***	0.2318	0.0033		0.0014	-0.1162	***	0.0085
Number of Observations	159, 231			159,231			159,231		
R-Square Adjusted				0.6102			0.6286		

*** indicates p-value<0.001, ** indicates p-value<0.01, * indicates p-value <0.1

Table (6)- Part of the Results of Model (4), (5), and (6)

Internal Rate of Return (IRR)

The model for this analysis is the same as model (4) and again we refer to table (6) to present its results. As this table illustrates, the adjusted R^2 is 0.61, and the posted interest rate, the estimated loss rate, and most of levels of loan grade and risk scores have significant relationship with IRR. Our findings show that IRR has an indirect relationship with IRR and increasing the posted interest rate would decrease IRR. However, more surprisingly, the estimated loss rate has a direct relationship with IRR, i.e., an increased the estimated loss rate (towards more losses) would increase the internal rate of return.

In addition, loan applications with grade “A”, “B”, “C”, and “D” would expect lower IRR, while loan applications with grade “E” and “HR” would expect higher IRR compared to the loan applications with grade “AA” (U-shaped pattern). Moreover, in accordance with our findings on the estimated loss score, but in contrast to natural expectation, and increased risk score (toward less riskier loans) almost everywhere correlates to a lower IRR. Finally, the findings show that there is no statistically significant relationship between high values of risk score (least risky) and IRR.

Return on Investment (ROI)

The model for this analysis is the same as model (6) and again we refer to table (6) to illustrate its results. The findings reveal that the adjusted $R^2$of the model is 0.63 and the posted interest rate, the estimated loss rate, few levels of loan grades, and all levels of risk score have statistically significant relationship with ROI. Following expectation, as the posted interest rate increases and the estimated loss rate decreases, ROI increases.

In addition, few grade levels have a significant relationship with ROI and in particular, the grade “HR” would quite surprisingly result in higher ROI compared to “AA”. Finally, while our results show that all levels of the risk score have statistically significant relationship with ROI, the order and values of its coefficient in our model are not consistent with the expectation and do not represent a robust pattern.

Discussion and Conclusion

In this paper, we presented a series of empirical models to study the performance of players in the P2P lending market. We have employed different sets of models to assess the interplay between the players in these markets, showing how signals from one class of participant effects the behavior of others in the market and may predict the loan outcomes.

Considering results of all our models, we find that the posted interest rate (as a platform signal), the investor responses, and the loan ex post outcome do not always align naturally. For example, stating “Business” as the loan purpose would result in a higher posted interest rate, lower percentage of loan funded, and longer funding duration compared to most of the other loan purposes including “Debt Consolidation” (table (5) and (6)). However, analyses of loan ex post outcomes reveal that the performance of “Business” loans, based on delinquency ratio and ROI are not statistically different from their peers in other groups. The only statistically dissimilar outcome of “Business” loans is their IRR, which showed the smallest indirect relationship compared to other loan purposes with significant relationship with IRR. Further, we expressed theoretical doubts about the use of IRR as an ex post aggregate measure (in contrast with a previous study), making this relationship between the “Business” purpose and its higher interest rate even more questionable, or perhaps explaining a possible source of confusion if the platform itself relies on the dubious IRR metric when setting a posted rate.

A similar inconsistency between platform decisions and ex post loan outcomes is found in the “Debt Consolidation” category. Actually if the P2P lending platform was mainly relying on either delinquency ratio or IRR for defining the posted interest rates, the “Debt Consolidation” should have one of the highest posted interest rates. However, loans in this category have the lowest posted interest rate among all loan purposes. Accordingly, the results show that in general, the difference of posted interest rates for various loan purposes cannot be explained by their expected outcomes and there is inconsistency between ex post outcome of loan applications and the platform signal for self-reported loan purposes.

The inconsistency between platform signals, investor response, and loan ex post outcomes is not limited to the posted interest rate as a function of loan purpose. This study expands the analysis of the relationship among the platform signals, investor responses, and loan ex post outcomes, and assesses the impact of four main platform signals, the posted interest rate, the estimated loss rate, the loan grade, and the risk score on the market.

Firstly, our analysis illustrates that an increased posted interest rate would attract investors in the fractional channel not only in terms of higher funded percentage of loan amount, but also in shorter funding duration. While the decision of investors in the fractional channel is in accordance with the expected ROI of loan applications, which increases as the posted interest rate increases, their decision is against their benefits based on the delinquency rate and IRR, since increasing the posted interest rate would increase the delinquency ratio and decrease the IRR. In contrast, investors in the whole channel show less desire to invest in loan applications with a higher posted interest rate, which provide more benefit for them based on delinquency and IRR.

Secondly, our analysis indicates that the estimated loss rate, the next platform signal, is an effective predictor of delinquency and ROI, and investors in the fractional channel acknowledge this result. Indeed, a higher estimated loss rate would increase the possibility of delinquency and reduce ROI. At the same time, and a higher estimated loss rate decreases the interest of investors in the fractional channel, based on both the funded percentage of loan amount and the funding duration. However, unexpectedly, a higher estimated loss rate increases IRR; an effect that is acknowledged by investors in the whole channel who show more interest in loan applications with a higher estimated loss rate.

The third platform signal that we have assessed is the loan grade. The results of our models indicate that this signal has no statistically significant relationship with possibility of delinquency. However, for a few grade levels like “B” and “HR”, there is a significant relationship between the grade level, and both IRR and ROI. Based on IRR and ROI, investing in loan applications with “HR” grade is beneficial for investors while investing in loan applications with “B” grade is not beneficial. Nevertheless, investors in the fractional and whole channels significantly ignore this result and prefer loan applications with “B” grade to “HR” grade. Moreover, there is a consistent and robust inverse U-shaped response toward loan grade among investors in the fractional channel that shows more desire for middle grade loans among fractional investors. However, their desire is not statistically supported by their benefits, measured by delinquency, IRR, or ROI. There is also a statistically significant relationship between loan grade and investor desires in the whole channel. However, this significant, but not robust relationship cannot be supported by the outcomes of loans, measured by delinquency, IRR, or ROI.

The last platform signal, the risk score, shows a direct and strictly robust relationship with the delinquency ratio and IRR. Indeed, higher risk scores predict a lower delinquency ratio, but also lower IRR. This result is clearly acknowledged by investors in the fractional channel, where funding duration increases as the risk score increases toward better applications. In fact, investors in the fractional channel have a tendency to invest in loan applications with lower risk score (the riskier applications), which would be expected to lead to higher delinquency but more IRR. In contrast, investors in the whole channel are indifferent about risk score; their responses show no statistically significant relationship with the risk score.

In summary, not all platform signals are reliable predictors of loan application outcomes. In the studied market, the risk score is the only robust and consistent signal for prediction of loan delinquency in the market. Moreover, if IRR is the primary concern, one may find the loan grade as a significant predictor of IRR; however, the statistically significant U-shaped relationship between loan grade and IRR does not follow the anticipated increasing trend for benefit of different levels of loan grade, and the loan grades in their current form seem to be misinforming investors. In addition, if ROI is the primary concern, neither the loan grade nor the risk score are reliable predictor of ROI. Taking all these results together, we conclude that the studied P2P lending platform needs more consistent and robust signals about expected outcomes of loan applications to help investors, especially in the fractional market, with better metrics to analyze and compare loan applications.

References:

Adams, W., Einav, L., & Levin, J. (2009). Liquidity constraints and imperfect information in subprime lending. American Economic Review, 99(1), 49-84.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. Journal of the operational research society, 54(6), 627-635.

Balyuk, T. (2016). Financial Innovation and Borrowers: Evidence from Peer-to-Peer Lending.

Collier, B. C., & Hampshire, R. (2010, February). Sending mixed signals: Multilevel reputation effects in peer-to-peer lending markets. In Proceedings of the 2010 ACM conference on Computer supported cooperative work (pp. 197-206). ACM.

Cox, N. (2017). The Impact of Risk-Based Pricing in the Student Loan Market: Evidence from Borrower Repayment Decisions. SIEPR Working Paper.

De Roure, C., Pelizzon, L., & Tasca, P. (2016). How does P2P lending fit into the consumer credit market?.

Dietrich, A., & Wernli, R. (2016). What drives the interest rates in the p2p consumer lending market? empirical evidence from switzerland.

Edelberg, W. (2006). Risk-based pricing of interest rates for consumer loans. Journal of monetary Economics, 53(8), 2283-2298.

Emekter, R., Tu, Y., Jirasakuldech, B., & Lu, M. (2015). Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. Applied Economics, 47(1), 54-70.

Freedman, S. M., & Jin, G. Z. (2011). Learning by Doing with Asymmetric Information: evidence from Prosper. com (No. w16855). National Bureau of Economic Research.

Gonzalez, L., & Loureiro, Y. K. (2014). When can a photo increase credit? The impact of lender and borrower profiles on online peer-to-peer loans. Journal of Behavioral and Experimental Finance, 2, 44-58.

Herzenstein, M., Andrews, R. L., Dholakia, U. M., & Lyandres, E. (2008). The democratization of personal consumer loans? Determinants of success in online peer-to-peer lending communities. Boston University School of Management Research Paper, 14(6).

Herzenstein, M., Andrews, R. L., Dholakia, U. M., & Lyandres, E. (2008). The democratization of personal consumer loans? Determinants of success in online peer-to-peer lending communities. Boston University School of Management Research Paper, 14(6).

Komarova Loureiro, Y., & Gonzalez, L. (2015). Competition against common sense: insights on peer-to-peer lending as a tool to allay financial exclusion. International Journal of Bank Marketing, 33(5), 605-623.

Komarova Loureiro, Y., & Gonzalez, L. (2015). Competition against common sense: insights on peer-to-peer lending as a tool to allay financial exclusion. International Journal of Bank Marketing, 33(5), 605-623.

Lee, E., & Lee, B. (2012). Herding behavior in online P2P lending: An empirical investigation. Electronic Commerce Research and Applications, 11(5), 495-503.

Lin, M., Prabhala, N. R., & Viswanathan, S. (2009). Judging borrowers by the company they keep: Social networks and adverse selection in online peer-to-peer lending. Ssrn Elibrary.

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. Management Science, 59(1), 17-35.

Loureiro, Y. K., & Gonzalez, L. (2015). Competition against common sense: insights on peer-to-peer lending as a tool to allay financial exclusion. International Journal of Bank Marketing, 33(5), 605-623.

Mach, T., Carter, C., & Slattery, C. (2014). Peer-to-peer lending to small businesses.
Mach, T., Carter, C., & Slattery, C. (2014). Peer-to-peer lending to small businesses.
Miller, S. (2015). Information and default in consumer credit markets: Evidence from a natural experiment. Journal of Financial Intermediation, 24(1), 45-70.

Mitra, T., & Gilbert, E. (2014, February). The language that gets people to give: Phrases that predict success on kickstarter. In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing (pp. 49-61). ACM.

Nowak, A., Ross, A., & Yencha, C. (2018). Small Business Borrowing and Peer-To-Peer Lending: Evidence from Lending Club. Contemporary Economic Policy, 36(2), 318-336.
Nowak, A., Ross, A., & Yencha, C. (2018). Small Business Borrowing and Peer-To-Peer Lending: Evidence from Lending Club. Contemporary Economic Policy, 36(2), 318-336.
Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. Decision Support Systems, 89, 113-122.

Serrano-Cinca, C., Gutierrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. PloS one, 10(10), e0139427.

Singh, H., Gopal, R., & Li, X. (2008). Risk and Return of Investments in Online Peer-to-Peer Lending. University of Texas, 1-6.

Thomas, L. C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. International journal of forecasting, 16(2), 149-172.

Walke, A. G., Fullerton Jr, T. M., & Tokle, R. J. (2018). Risk-based loan pricing consequences for credit unions. Journal of Empirical Finance, 47, 105-119.

West, D. (2000). Neural network credit scoring models. Computers & Operations Research, 27(11-12), 1131-1152.

Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. Journal of Financial and Quantitative Analysis, 15(3), 757-770.

Xin, Y. (2018). Asymmetric Information, Reputation, and Welfare in Online Credit Markets.

Zhang, J., & Liu, P. (2012). Rational herding in microloan markets. Management science, 58(5), 892-912.

Appendix

		Loan application Duration (hour per $1000)													
				Investment Channel						Default					
		Posted Interest Rate		Fractional		Whole		percent_funded		0	1	IRR		ROI	
Grade	**Column %**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Row %**	**Row %**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
HR	4.97%	0.313	0.010	33.148	39.986	0.327	1.333	0.779	0.382	68.77%	31.23%	-0.037	0.167	0.149	0.461
E	7.82%	0.274	0.025	11.072	23.374	0.251	1.990	0.940	0.224	71.56%	28.44%	-0.025	0.137	0.128	0.404
D	11.34%	0.224	0.023	10.722	19.413	0.153	0.755	0.930	0.239	74.55%	25.45%	-0.021	0.123	0.102	0.357
C	17.09%	0.173	0.019	7.015	29.569	0.153	2.276	0.967	0.171	78.51%	21.49%	-0.014	0.098	0.081	0.302
B	20.69%	0.134	0.015	8.475	16.758	0.147	2.620	0.969	0.160	83.30%	16.70%	-0.009	0.082	0.069	0.250
A	25.45%	0.101	0.013	9.807	15.692	0.191	10.838	0.970	0.155	88.46%	11.54%	-0.004	0.060	0.063	0.193
AA	12.64%	0.071	0.007	9.069	43.746	0.153	4.253	0.966	0.162	93.91%	6.09%	-0.000	0.045	0.054	0.137
Risk Score															
1	0.43%	0.312	0.014	32.352	40.098	0.001	.	0.614	0.460	43.33%	56.67%	-0.085	0.232	0.003	0.557
2	5.91%	0.254	0.046	10.506	30.979	0.224	1.881	0.949	0.212	72.52%	27.48%	-0.026	0.136	0.117	0.395
3	7.98%	0.227	0.055	14.511	31.781	0.191	1.788	0.947	0.213	74.75%	25.25%	-0.021	0.122	0.104	0.360
4	15.65%	0.193	0.060	14.582	29.211	0.176	2.928	0.957	0.191	76.66%	23.34%	-0.018	0.108	0.083	0.329
5	9.94%	0.184	0.067	14.613	30.279	0.143	0.739	0.944	0.214	78.87%	21.13%	-0.015	0.107	0.087	0.313
6	11.85%	0.160	0.057	11.726	19.813	0.160	2.912	0.947	0.209	81.55%	18.45%	-0.013	0.097	0.078	0.286
7	11.70%	0.136	0.047	10.668	29.077	0.306	17.435	0.958	0.185	84.32%	15.68%	-0.008	0.080	0.073	0.245
8	15.19%	0.114	0.034	12.492	18.164	0.137	1.008	0.952	0.196	87.36%	12.64%	-0.005	0.068	0.068	0.214
9	9.31%	0.098	0.025	11.611	18.583	0.124	0.872	0.960	0.176	90.49%	9.51%	-0.002	0.057	0.065	0.179
10	5.71%	0.088	0.023	9.675	58.389	0.108	0.450	0.952	0.194	92.94%	7.06%	-0.000	0.049	0.065	0.150
11	6.34%	0.077	0.019	6.740	10.556	0.122	0.624	0.980	0.123	95.07%	4.93%	0.001	0.037	0.062	0.123

Table (A.1)- Descriptive analysis of the grade and risk score

		Loan application Duration (hour per $1000)													
				Investment Channel						Default					
		Posted Interest Rate		Fractional		Whole		percent_funded		0	1	IRR		ROI	
Loan Purpose	**Column %**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Row %**	**Row %**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
Auto/Motorcycle/RV/Boat	2.02%	0.182	0.083	18.186	48.812	0.223	1.045	0.885	0.299	83.64%	16.36%	-0.006	0.090	0.121	0.293
Baby & Adoption	0.17%	0.169	0.082	11.922	22.251	0.247	1.149	0.912	0.268	80.20%	19.80%	-0.020	0.120	0.070	0.305
Business	3.92%	0.190	0.081	18.399	29.372	0.145	2.052	0.884	0.289	72.68%	27.32%	-0.029	0.132	0.057	0.370
Debt Consolidation	70.78%	0.147	0.064	9.451	20.379	0.152	5.759	0.968	0.165	83.06%	16.94%	-0.010	0.087	0.077	0.267
Green Loans	0.07%	0.171	0.079	18.926	28.894	0.230	0.902	0.918	0.261	77.78%	22.22%	-0.034	0.134	0.052	0.378
Household Expenses	9.56%	0.168	0.078	15.403	29.171	0.317	13.269	0.925	0.242	81.76%	18.24%	-0.015	0.110	0.093	0.307
Medical / Dental	2.34%	0.159	0.075	12.674	24.311	0.234	0.906	0.960	0.181	82.74%	17.26%	-0.015	0.113	0.084	0.305
Not Available	0.08%	0.150	0.070	30.904	35.942	0.136	0.316	0.900	0.281	83.76%	16.24%	-0.008	0.067	0.078	0.274
Other	8.48%	0.172	0.083	19.664	48.343	0.179	0.849	0.908	0.268	81.35%	18.65%	-0.017	0.113	0.081	0.314
Taxes	0.82%	0.173	0.075	10.743	21.053	0.181	0.704	0.941	0.221	84.93%	15.07%	-0.005	0.095	0.136	0.275
Vacation/Special Occasion	0.98%	0.162	0.079	13.931	25.525	0.280	0.941	0.946	0.210	80.53%	19.47%	-0.013	0.104	0.094	0.317
Wedding Loans	0.79%	0.179	0.075	10.729	19.681	0.126	0.577	0.919	0.255	86.28%	13.72%	-0.004	0.087	0.134	0.278

Table (A.2) - Descriptive analysis of the loan purpose

		Loan application Duration (hour per $1000)											
		Investment Channel						Default					
		Fractional		Whole		percent_funded		0	1	IRR		ROI	
Round of Posted Interest Rate	Column %	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Row %	Row %	Mean	Std Dev	Mean	Std Dev
5	0.02%	12.591	12.714	.	.	0.805	0.354	100.00%	0.00%	0.005	.	0.071	.
6	5.65%	10.146	68.048	0.161	1.466	0.976	0.134	95.11%	4.89%	0.001	0.037	0.051	0.120
7	6.58%	8.795	19.960	0.306	17.778	0.974	0.145	92.04%	7.96%	-0.002	0.055	0.050	0.158
8	6.66%	9.877	10.025	0.250	13.292	0.964	0.166	90.46%	9.54%	-0.002	0.050	0.055	0.168
9	5.66%	9.903	16.170	0.129	1.159	0.965	0.166	89.14%	10.86%	-0.003	0.054	0.059	0.188
10	8.10%	8.760	11.380	0.135	0.872	0.981	0.128	86.41%	13.59%	-0.006	0.069	0.058	0.211
11	8.15%	7.583	20.407	0.148	3.516	0.983	0.118	85.74%	14.26%	-0.006	0.071	0.066	0.222
12	5.30%	6.279	10.814	0.134	1.361	0.973	0.152	84.77%	15.23%	-0.008	0.076	0.068	0.236
13	6.62%	7.251	12.047	0.161	3.443	0.979	0.135	82.24%	17.76%	-0.012	0.089	0.062	0.260
14	5.38%	7.715	14.406	0.117	0.566	0.967	0.169	81.96%	18.04%	-0.010	0.086	0.075	0.264
15	4.36%	6.130	14.790	0.143	0.731	0.973	0.151	80.64%	19.36%	-0.010	0.085	0.080	0.274
16	3.97%	6.892	22.222	0.136	1.704	0.961	0.182	79.38%	20.62%	-0.013	0.094	0.079	0.292
17	4.02%	10.857	41.570	0.142	0.844	0.953	0.198	78.60%	21.40%	-0.015	0.102	0.083	0.306
18	2.42%	3.651	11.819	0.160	1.057	0.970	0.164	78.28%	21.72%	-0.012	0.093	0.097	0.304
19	3.23%	6.229	27.697	0.150	0.646	0.977	0.145	78.24%	21.76%	-0.014	0.103	0.102	0.312
20	2.79%	9.967	28.677	0.148	0.685	0.929	0.239	74.91%	25.09%	-0.023	0.124	0.084	0.348
21	2.72%	8.983	18.369	0.124	0.479	0.940	0.223	75.60%	24.40%	-0.019	0.115	0.101	0.346
22	1.37%	2.942	7.535	0.139	1.023	0.967	0.174	72.28%	27.72%	-0.023	0.124	0.097	0.360
23	1.62%	7.381	17.380	0.213	1.296	0.938	0.230	75.16%	24.84%	-0.016	0.111	0.121	0.356
24	2.16%	6.520	19.959	0.247	1.845	0.947	0.211	73.21%	26.79%	-0.021	0.125	0.117	0.371
25	2.05%	15.077	19.753	0.328	4.054	0.863	0.318	73.71%	26.29%	-0.027	0.142	0.126	0.395
26	2.16%	14.433	27.056	0.214	0.796	0.906	0.268	71.98%	28.02%	-0.024	0.132	0.126	0.395
27	0.74%	4.971	24.944	0.190	0.751	0.953	0.208	74.30%	25.70%	-0.020	0.133	0.169	0.393
28	1.41%	4.563	14.261	0.269	0.945	0.971	0.166	73.51%	26.49%	-0.022	0.135	0.148	0.397
29	1.17%	4.716	11.661	0.256	1.130	0.942	0.229	72.34%	27.66%	-0.027	0.153	0.169	0.417
30	0.65%	7.687	16.540	0.368	1.324	0.905	0.287	67.49%	32.51%	-0.043	0.169	0.117	0.469
31	5.00%	32.960	37.773	0.218	1.583	0.738	0.399	67.40%	32.60%	-0.038	0.170	0.147	0.474
32	0.01%	3.279	3.323	.	.	1.000	0.000	63.16%	36.84%	-0.001	0.062	0.209	0.387

Table (A.3) – Descriptive analysis of the round of posted interest rates.

	Poster Interest Rate (Model 1)			Loan Funded Percentage (Model 2)			Funding duration (Model 3)					
							Fractional			Whole		
Loan Purpose	Coeff.		Std. Error	Coeff.		Std. Error	Coeff.		Std. Error	Coeff.		Std. Error
Auto/Motorcycle/RV/Boat	-0.004	***	0.001	-0.002		0.009	1.963		4.357	5.985	***	0.689
Baby & Adoption	0.009	***	0.002	-0.040		0.020	10.244		9.619	-0.365		2.125
Business	0.005	***	0.001	-0.043	***	0.008	27.076	***	4.052	9.535	***	0.607
Debt Consolidation	-0.008	***	0.001	0.000		0.008	9.399	*	3.774	6.324	***	0.540
Green Loans	0.008	*	0.003	0.024		0.033	22.264		16.398	-7.702	**	2.833
Household Expenses	0.000		0.001	-0.023	**	0.008	16.387	***	3.891	7.532	***	0.558
Medical / Dental	0.005	***	0.001	-0.015		0.010	12.110	**	4.585	1.144	*	0.632
Not Available	-0.004		0.003	-0.197	***	0.035	96.922	***	16.900	-7.022	**	2.150
Other	-0.001		0.001	-0.020	*	0.008	21.699	***	3.933	3.543	***	0.565
Vacation/Special Occasion	-0.001		0.001	-0.012		0.011	14.263	**	5.341	2.062	**	0.778
Wedding Loans	0.004	**	0.001	-0.004		0.011	9.202	*	5.340	5.318	***	0.869
P2P Market Signals												
Posted Interest Rate				1.798	***	0.090	-164.156	***	40.601	68.630	***	7.296
Estimated Loss Rate				-0.299	***	0.010	59.128	***	3.496	-1.335	**	0.482
Grade												
GRADE="A"				0.154	***	0.008	-50.555	***	2.666	3.602	***	0.269
GRADE="B"				0.254	***	0.011	-97.003	***	4.019	1.433	***	0.378
GRADE="C"				0.294	***	0.014	-118.018	***	5.104	2.100	***	0.481

GRADE="D"				0.304	***	0.016	-116.676	***	6.284	0.303		0.626
GRADE="E"				0.300	***	0.019	-115.748	***	7.685	1.820	*	0.824
GRADE="HR"				0.180	***	0.021	-56.583	***	8.352	1.739	*	1.040
Risk Score												
RISK_SCORE=2				0.105	***	0.006	-9.291	**	3.543	14.519		20972.250
RISK_SCORE=3				0.095	***	0.006	1.889		3.503	15.742		20972.250
RISK_SCORE=4				0.103	***	0.006	6.415	*	3.430	19.729		20972.250
RISK_SCORE=5				0.120	***	0.006	1.953		3.502	26.815		20972.250
RISK_SCORE=6				0.099	***	0.006	14.182	***	3.588	22.708		20972.250
RISK_SCORE=7				0.103	***	0.007	18.523	***	3.713	23.118		20972.250
RISK_SCORE=8				0.105	***	0.007	33.462	***	3.791	23.576		20972.250
RISK_SCORE=9				0.110	***	0.008	31.225	***	4.055	25.641		20972.250
RISK_SCORE=10				0.096	***	0.010	17.189	***	4.375	25.759		20972.250
RISK_SCORE=11				0.051	***	0.011	27.562	***	4.613	35.137		20972.250
Intercept	0.143	***	0.005	-0.309	***	0.061	97.433	***	3.765	76.038		20973.270
Control Variables												
Total Inquiries	0.001	***	0.000	0.002	***	0.000	-0.686	***	0.135	-0.038	**	0.014
Total Open Revolving Accounts	0.000	**	0.000	0.001	***	0.000	-0.407	***	0.104	0.025	*	0.012
Revolving Balance	0.000	***	0.000	0.000		0.000	0.000		0.000	0.000		0.000
Revolving Available Percent	0.000	***	0.000	0.000	*	0.000	0.048	*	0.026	-0.004		0.004
Loan Application Amount	0.000	***	0.000	0.000	***	0.000	0.007	***	0.000	0.000	***	0.000
Inflation Rate	0.006	***	0.000	-0.025	***	0.002	24.620	***	1.155	-22.335	***	0.162
Credit History Length	0.000	***	0.000	-0.001	***	0.000	0.510	***	0.055	-0.023	***	0.006
Bankcard Utilization	0.006	***	0.001	0.035	***	0.004	-10.332	***	2.201	2.621	***	0.332
Debt to income with the loan	0.000	***	0.000	0.000	***	0.000	0.000	***	0.000	0.000	***	0.000
Delinquencies within last7 years	0.000	***	0.000	0.000	***	0.000	0.559	***	0.048	0.194	***	0.005
Delinquencies over 30 days	0.000	***	0.000	0.000	***	0.000	-0.106	*	0.064	-0.153	***	0.009
Inquiries last 6 months	0.008	***	0.000	-0.007	***	0.001	3.236	***	0.346	0.213	***	0.041
Federal Reserve Rate	0.086	***	0.006	-0.905	***	0.048	705.043	***	22.714	-775.958	***	4.175
FICO_SCO="< 600"	0.081	***	0.004	0.044		0.036	-8.426		15.933	35.351		141.681
FICO_SCO="600-619"	0.075	***	0.003	0.046		0.030	-17.160		11.944	22.889		150.432
FICO_SCO="620-639"	0.076	***	0.002	0.055	*	0.027	-16.455	*	9.101	17.072		49.522
FICO_SCO="640-659"	0.082	***	0.002	0.061	*	0.024	-21.964	**	7.068	5.501	***	0.862
FICO_SCO="660-679"	0.051	***	0.002	0.063	**	0.024	-18.153	**	6.976	5.795	***	0.848
FICO_SCO="680-699"	0.038	***	0.002	0.063	**	0.024	-17.191	*	6.948	8.242	***	0.840
FICO_SCO="700-719"	0.031	***	0.002	0.073	**	0.024	-24.659	***	6.943	8.591	***	0.837
FICO_SCO="720-739"	0.026	***	0.002	0.079	**	0.024	-26.630	***	6.947	9.760	***	0.832
FICO_SCO="740-759"	0.021	***	0.002	0.076	**	0.024	-23.667	***	7.063	8.056	***	0.837
FICO_SCO="760-779"	0.017	***	0.002	0.066	**	0.025	-23.008	**	7.190	9.528	***	0.844
FICO_SCO="780-799"	0.010	***	0.002	0.057	*	0.026	-17.281	*	7.469	3.362	***	0.882
FICO_SCO="800-819"	0.003		0.002	0.045		0.028	-12.472		7.951	11.243	***	0.898

FICO_SCO="NA"	0.060	***	0.002	0.046	*	0.025	19.723	*	7.819			
SCOREX="< 600"	0.061	***	0.001	-0.010		0.011	-4.698		5.462	10.042	***	0.441
SCOREX="600-619"	0.057	***	0.001	0.029	***	0.007	-29.247	***	3.550	2.508	***	0.382
SCOREX="620-639"	0.052	***	0.001	0.022	***	0.006	-24.642	***	2.938	-0.415		0.326
SCOREX="640-649"	0.049	***	0.001	-0.041	***	0.006	1.462		2.898	11.187	***	0.325
SCOREX="650-664"	0.041	***	0.001	-0.036	***	0.006	9.271	***	2.450	4.159	***	0.292
SCOREX="665-689"	0.041	***	0.000	-0.011	*	0.005	3.784	*	2.143	1.389	***	0.264
SCOREX="690-701"	0.035	***	0.001	0.000		0.005	1.758		2.276	-0.946	***	0.276
SCOREX="702-723"	0.030	***	0.000	0.006		0.005	-0.645		2.041	-0.383		0.240
SCOREX="724-747"	0.021	***	0.000	0.000		0.005	2.140		1.926	-0.059		0.230
SCOREX="748-777"	0.014	***	0.000	0.005		0.005	-1.346		1.843	-1.801	***	0.220
SCOREX="NA"	0.014		0.008	-0.056		0.044	3.281		21.108			
Loan Origination Year =2011	0.024	***	0.001	-0.060	***	0.009	30.428	***	4.171			
Loan Origination Year=2012	0.020	***	0.001	0.005		0.008	-17.302	***	3.824			
Loan Origination Year=2014	-0.038	***	0.000	0.018	***	0.003	-0.024		1.480	22.915	***	3.428
Loan Origination Year=2015	-0.043	***	0.000	0.026	***	0.006	34.724	***	2.427	12.665	***	3.439
Homeownership="True"	0.009	***	0.000	0.015	***	0.002	-7.781	***	0.978	1.900	***	0.114
Income_Range="$100,000+"	-0.040	***	0.005	-0.050	*	0.030	39.080	*	16.571	-21.553		206.786
Income_Range="$1-24,999"	-0.006		0.005	-0.086	**	0.030	62.916	***	16.604	-23.736		206.786
Income_Range="$25,000-49,999"	-0.022	***	0.005	-0.062	*	0.030	50.081	***	16.548	-23.923		206.786
Income_Range="$50,000-74,999"	-0.029	***	0.005	-0.055	*	0.030	42.320	*	16.552	-22.070		206.786
Income_Range="$75,000-99,999"	-0.037	***	0.005	-0.053	*	0.030	42.382	*	16.568	-21.383		206.786
Income_Range="Not employed"	-0.004		0.005	-0.138	***	0.031	95.901	***	16.904	-13.343		222.064
Number of Observations	284,983			50,709			50,709			137,641		
R-Square Adjusted	0.6109			0.1912			0.3435			0.3585		

*** indicates p-value<0.001, ** indicates p-value<0.01, * indicates p-value <0.1

Table (A.4) - Results of models (1), (2), and (3)

	Default (Model 4)			IRR (Model 5)			ROI (Model 6)		
	Coefficient		**Std. Error**	**Coefficient**		**Std. Error**	**Coefficient**		**Std. Error**
Auto/Motorcycle/RV/Boat	-0.0781		0.1029	-0.0008	*	0.0005	-0.0035		0.0031
Baby & Adoption	0.4104	*	0.2038	-0.0006		0.0011	-0.0082		0.0072
Business	0.5306	***	0.0913	-0.0010	*	0.0004	0.0029		0.0028
Debt Consolidation	0.1333		0.0848	-0.0002		0.0004	-0.0006		0.0025
Green Loans	0.2814		0.3254	-0.0009		0.0018	-0.0109		0.0117
Household Expenses	0.2050	*	0.0875	-0.0005		0.0004	0.0031		0.0026
Medical / Dental	0.2099	*	0.0971	0.0002		0.0005	0.0043		0.0030
Not Available	0.1476	*	0.2720	-0.0004		0.0014	0.0052		0.0093
Other	0.2107	*	0.0876	-0.0009	*	0.0004	0.0064	*	0.0026
Vacation/Special Occasion	0.2171	*	0.1156	-0.0004		0.0006	0.0000		0.0037
Wedding Loans	-0.1939		0.1227	-0.0003		0.0005	-0.0032		0.0036
P2P Lending Market Signals									
Posted Interest Rate	2.4334	**	0.9089	-0.0351	***	0.0045	1.6143	***	0.0305

Estimated Loss Rate	0.2279	**	0.0855	0.0082	***	0.0004	-0.0124	***	0.0028
Grade									
GRADE="A"	0.0243		0.0534	-0.0016	***	0.0003	0.0027		0.0018
GRADE="B"	0.0242		0.0736	-0.0034	***	0.0004	0.0046	*	0.0025
GRADE="C"	0.0486		0.0894	-0.0023	***	0.0005	0.0048		0.0031
GRADE="D"	0.0398		0.1053	-0.0009		0.0006	0.0028		0.0037
GRADE="E"	0.0628		0.1236	0.0027	***	0.0007	0.0035		0.0044
GRADE="HR"	0.0378		0.1381	0.0040	***	0.0007	0.0137	**	0.0048
Risk Score									
RISK_SCORE=2	-1.0870	***	0.2214	0.0055	***	0.0014	-0.1159	***	0.0082
RISK_SCORE=3	-1.1247	***	0.2212	0.0051	***	0.0014	-0.1183	***	0.0082
RISK_SCORE=4	-1.1326	***	0.2208	0.0044	**	0.0013	-0.1166	***	0.0082
RISK_SCORE=5	-1.1887	***	0.2212	0.0045	***	0.0014	-0.1184	***	0.0082
RISK_SCORE=6	-1.2278	***	0.2216	0.0037	**	0.0014	-0.1141	***	0.0082
RISK_SCORE=7	-1.2752	***	0.2221	0.0035	**	0.0014	-0.1169	***	0.0083
RISK_SCORE=8	-1.3689	***	0.2225	0.0043	**	0.0014	-0.1173	***	0.0083
RISK_SCORE=9	-1.5018	***	0.2240	0.0019		0.0014	-0.1173	***	0.0083
RISK_SCORE=10	-1.6273	***	0.2273	0.0019		0.0014	-0.1167	***	0.0084
RISK_SCORE=11	-1.8287	***	0.2318	0.0033		0.0014	-0.1162	***	0.0085
Intercept	-0.7733		0.7696	0.0446	***	0.0045	0.0091		0.0297
Control Variables									
Total Inquiries	0.0094	***	0.0020	0.0000		0.0000	-0.0011	***	0.0001
Total Open Revolving Accounts	0.0131	***	0.0017	0.0001	***	0.0000	0.0002	***	0.0001
Revolving Balance	0.0000	***	0.0000	0.0000		0.0000	0.0000	*	0.0000
Revolving Available Percent	0.0004		0.0005	0.0000		0.0000	-0.0002	***	0.0000
Loan Application Amount	0.0000	***	0.0000	0.0000	***	0.0000	0.0000	*	0.0000
Inflation Rate	-0.0280		0.0229	0.0006	***	0.0001	0.0020	**	0.0008
Credit History Length	0.0028	**	0.0009	0.0000	***	0.0000	0.0003	***	0.0000
Bankcard Utilization	-0.3028	***	0.0449	-0.0012	***	0.0002	-0.0096	***	0.0015
Debt to income with the loan	0.0000	***	0.0000	0.0000	***	0.0000	0.0000	***	0.0000
Delinquencies within last7 years	-0.0052	***	0.0008	0.0000		0.0000	-0.0002	***	0.0000
Delinquencies over 30 days	-0.0020	*	0.0012	0.0000	*	0.0000	0.0005	***	0.0000
Inquiries last 6 months	0.0215	***	0.0057	-0.0001	**	0.0000	-0.0007	**	0.0002
Federal Reserve Rate	1.8798	***	0.5459	-0.0132	***	0.0027	0.0502	**	0.0178
FICO_SCO="< 600"	1.7211	***	0.4390	0.0018		0.0025	-0.0028		0.0165
FICO_SCO="600-619"	0.7731	*	0.3723	0.0022		0.0016	0.0004		0.0109
FICO_SCO="620-639"	0.9049	***	0.2536	0.0022	*	0.0011	-0.0025		0.0075
FICO_SCO="640-659"	0.5216	**	0.1998	0.0016	*	0.0008	-0.0089		0.0057
FICO_SCO="660-679"	0.5271	**	0.1987	0.0008		0.0008	-0.0063		0.0057
FICO_SCO="680-699"	0.5531	**	0.1981	0.0004		0.0008	-0.0046		0.0057
FICO_SCO="700-719"	0.5024	*	0.1978	0.0001		0.0008	-0.0016		0.0057
FICO_SCO="720-739"	0.3826	*	0.1978	-0.0007		0.0008	-0.0033		0.0057
FICO_SCO="740-759"	0.3182		0.1990	-0.0009		0.0008	-0.0038		0.0057
FICO_SCO="760-779"	0.3097		0.2017	-0.0017	*	0.0008	-0.0036		0.0058
FICO_SCO="780-799"	0.1314		0.2096	-0.0012		0.0008	-0.0073		0.0060
FICO_SCO="800-819"	0.2268		0.2237	-0.0018	*	0.0009	-0.0094		0.0065
FICO_SCO="NA"	0.1919		0.2522	0.0015		0.0010	0.0005		0.0070

SCOREX="< 600"	1.0328	***	0.0649	-0.0045	***	0.0004	0.0092	***	0.0024
SCOREX="600-619"	0.9367	***	0.0577	-0.0046	***	0.0003	0.0038	*	0.0021
SCOREX="620-639"	0.7823	***	0.0506	-0.0045	***	0.0003	0.0041	*	0.0017
SCOREX="640-649"	0.7133	***	0.0510	-0.0046	***	0.0003	0.0031	*	0.0018
SCOREX="650-664"	0.5998	***	0.0466	-0.0046	***	0.0002	0.0058	***	0.0016
SCOREX="665-689"	0.5433	***	0.0427	-0.0044	***	0.0002	0.0036	*	0.0014
SCOREX="690-701"	0.4924	***	0.0445	-0.0043	***	0.0002	0.0029	**	0.0015
SCOREX="702-723"	0.4115	***	0.0409	-0.0045	***	0.0002	0.0024	*	0.0013
SCOREX="724-747"	0.3067	***	0.0400	-0.0044	***	0.0002	0.0030	*	0.0013
SCOREX="748-777"	0.1545	***	0.0401	-0.0034	***	0.0002	0.0014		0.0013
SCOREX="NA"	-0.1570		0.7826	-0.0073	*	0.0033	-0.0622	**	0.0240
Loan Origination Year =2011	0.8347	***	0.1604	0.0054	***	0.0007	-0.0087	*	0.0043
Loan Origination Year=2012	0.6381	***	0.1567	0.0019	**	0.0006	-0.0031		0.0041
Loan Origination Year=2014	0.2068	***	0.0290	-0.0017	***	0.0001	0.0002		0.0009
Loan Origination Year=2015	0.3057	***	0.0428	-0.0016	***	0.0002	0.0036	*	0.0014
Homeownership="True"	-0.0409	*	0.0167	0.0000		0.0001	0.0003		0.0006
Income_Range="$100,000+"	-1.5238	**	0.5492	0.0007		0.0035	0.0342		0.0237
Income_Range="$1-24,999"	-0.7438		0.5496	0.0021		0.0035	0.0485	*	0.0237
Income_Range="$25,000-49,999"	-1.0123	*	0.5490	0.0020		0.0035	0.0405	*	0.0237
Income_Range="$50,000-74,999"	-1.2039	*	0.5490	0.0010		0.0035	0.0347	*	0.0237
Income_Range="$75,000-99,999"	-1.3559	*	0.5492	0.0013		0.0035	0.0334		0.0237
Income_Range="Not employed"	-0.9322	*	0.5593	0.0004		0.0036	0.0319		0.0240
Number of Observations				159,231			159,231		
R-Square Adjusted				0.6102			0.6286		

*** indicates p-value<0.001, ** indicates p-value<0.01, * indicates p-value <0.1

Table (A.5) – Results of model (4), (5), and (6)

Chapter 5. Conclusions

In this thesis, we addressed three main questions. First, we explored the many-to-many matching problem, and in particular the course allocation problem and developed optimization-based mechanisms that improve fairness and efficiency and can reduce the benefit of students who misreport their preferences. Second, the team formation problem was presented as an extension of the roommate problem, and an algorithm for achieving game-theoretically stable and economically efficient solutions was developed. Third, we turned our attention to a peer-to-peer lending market and presented a series of empirical models to assess the interplay between players and show how signals from one class of participant predict the loan outcomes and affect the behavior of others in the market.

In detail, the first essay explored five new algorithmic variations to solve the course allocation problem, addressing the relative balance of efficiency, fairness, and incentive compatibility. To understand the performance of these five algorithms, we compared them to the draft and bidding point mechanisms, benchmarks that have been used in practice. We introduced a comprehensive systemization of natural metrics for these objectives, with a novel perspective of analyzing any set of outcomes based on the total (or average), range, and standard deviation of binary, ordinal, and cardinal utility. Our two-part OC algorithm (global ordinal optimization with cardinal tie-breaking) proved to be a very fair and efficient method, with alternative new methods (based on within-round optimization) performing better on incentive compatibility.

In the second paper, we investigated computational models for the team formation problem. In order to balance intra-team utility with solution stability, we formulated a bi-level binary optimization model and developed a branch-cut-and-price solution algorithm. The pseudo-polynomial approach to the bi-level problem is itself an interesting contribution, and we detailed how to implement the resulting algorithm, which still has to manage an exponential number of variables and constraints, which demanded the* development of advanced computational methods. Experimental results indicated that the proposed algorithm BCP is particularly effective at finding high-quality solutions quickly. Stability as an objective to be optimized over remains a particularly challenging computational problem, but we have shown a promising new approach. Our results also indicated that ignoring stability can result in inferior solutions when Pareto improvements are available.

In the third paper, having collected a large dataset of publicly available loan information for over four years of loan origination requests (with all follow-up data through the completion of 36-month loan terms) from an anonymous lending platform, we analyzed the interplay between the players in P2P lending markets. In particular, we first explored the borrowers' disclosed personal information and analyzed the response of the platform as well as the investors to these signals. Then, we analyzed the response of investors to the platform signals and examined how closely investors follow the signals provided by the platform. Finally, we studied the effectiveness of the P2P lending platform signals in predicting the success of loans. Our results suggest that certain platform signals might be misleading, as well as significant trends in which the self-reported loan purpose might also be a signal that is being misread by investors.

As a whole, these essays showed that matching markets, which might at first seem simple, are actually incredibly complex. Though we illustrate how difficult they may sometimes be to understand, predict, or control, we also demonstrated that advanced computational techniques create new opportunities for better understanding of existing markets and better designs for the markets of tomorrow.